IDENTIFICATION OF IRREVERSIBLE INHIBITORS OF DNA POLYMERASE β

by
Shelby Yuhas

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland
December 2020

© 2020 Shelby Yuhas
All rights reserved

Abstract

Many chemotherapeutics induce cell death by attacking deoxyribonucleic acid (DNA). Unfortunately, many cancer drugs suffer from poor selectivity and high toxicity. In addition to poor selectivity, the cytotoxic effects of DNA damaging agents are often reduced by DNA repair mechanisms. To combat this, DNA repair inhibition is an attractive route for enhancing the efficacy of cancer treatment, either by monotherapy or combination therapy. An established target for DNA repair inhibition is DNA polymerase beta (Pol β), the primary polymerase in base excision repair (BER). Pol β is an attractive target because it is overexpressed and often mutated in certain cancers. However, current Pol β inhibitors lack the potency and selectivity necessary for therapeutic relevance.

Small molecule libraries were synthesized and screened against WT Pol β and some select cancer-associated mutants. These compounds covalently and irreversibly inhibit Pol β. The relationship between domain binding and activity inactivation was explored using first-generation inhibitors that selectively bind to different Pol β domains. In addition, a potent (IC_{50} = 409-458 nM, K_I = 1.8 μM, k_{inact} = 7.0 s^{-1}) and selective second-generation inhibitor was identified against WT Pol β. This inhibitor selectively inactivated the polymerase activity via the disruption of DNA binding. A pro-inhibitor based on this molecule enhanced the cell killing effect of DNA damaging agents (i.e. methyl methanesulfonate (MMS) and bleomycin (BLM)) synergistically in HeLa and MCF-7 cells.

The screen against mutant forms of Pol β did not yield any potent or selective inhibitors. The mutant library ultimately failed due to a fault in the screening method and a path forward to overcome this shortcoming has been outlined. The lack of a mutant inhibitor was disappointing

because while the WT Pol β inhibitor enhanced the cytotoxic effects of DNA damaging agents, it did not address the concern of producing selectivity for cancer cells versus healthy cells.

DNA repair inhibition is also clinically useful for exploiting a cancer cell's inherent weakness by inducing synthetic lethality. Synthetic lethality is also attractive because it negates the need for DNA damaging agents. To establish selectivity for cancer cells, a synthetic lethal interaction between the Pol β inhibitor and an HR-deficiency was evaluated. The pro-inhibitor was synthetic lethal in BRCA1-deficient cells (MDA-MB-436). This effect was confirmed by knocking down Pol β by siRNA. This is the first report, to our knowledge, that confirms Pol β and BRCA1 synthetic lethality.

Advisor: Prof. Marc Greenberg

Readers: Prof. Rebekka Klausen

Prof. Steven Rokita

Acknowledgements

I am incredibly grateful to all the people who have helped me during my PhD work. Thank you to all my friends in Baltimore, at Johns Hopkins and at the Village church. Thanks to the former and current members of the Greenberg lab for their help, advice, and support over the past five years. Thank you to Dr. Daniel Laverty, Prof. Kun Yang, and Dr. Liwei Zheng for their encouragement, and willingness to help troubleshoot an experiment or reaction. Thank you to Marco Jacinto, Huijin Lee, Tingyu Wen, Alexandra Gittens, Haozhe Yang, and all other members, past and present, for their continued support and for creating a truly enjoyable environment in the lab. A special thanks to my friend and colleague, Jamie Alley, for her friendship, advice on experiments, and willingness to join me for a coffee break!

I am extremely thankful to my advisor, Prof. Greenberg, for his mentorship during my time in his lab. His guidance, questions, and advice has helped me grow into a better scientist. His dedication to his research and his students' success has and will continue to motivate me in my future endeavors.

Thank you to the readers of my thesis, Prof. Klausen and Prof. Rokita, for their support during my candidacy. I would like to thank the University staff, Ananya Majumdar for his assistance in collecting 800 Mz NMR data, Katie Tripp for her training and help with biophysical instruments, and Phil Mortimer for his support in the mass spectrometry facility.

Lastly, I am especially grateful for my family. Thank you to my parents and my brother, who have supported and visited me throughout this journey and have listened to me talk about my research for many hours. My husband, who moved to Baltimore and was incredibly understanding when I worked long days and weekends, continues to be an amazing support system and foundation for me. I am so thankful for his love and superior cooking skills.

Table of Contents

List of Tables

List of Figures

List of Schemes

List of Abbreviations

5IdU	5-iodo-deoxyuridine
5FU	5-fluorouracil
5FdUTP	5-fluoro-deoxyuridine triphosphate
8-oxodG	8-oxo-7,8-dihydro-2'-deoxyguanosine
A	fluorescence anisotropy
Ac	acetate
AP	apurinic/apyrimidinic
Ape1	AP endonuclease 1
ARP	aldehyde reaction probe
AZT	3'-azidothymidine
B-al	5'-nucleoside-5'-aldehydes
BER	base excision repair
BHQ	Black Hole Quencher
BLM	bleomycin
BRCT	BRCA1 C Terminus
BSA	bovine serum albumin
C	palm subdomain
C2-AP	C2'-oxidized abasic lesion
C4-AP	C4'-oxidized abasic lesion
CHO	Chinese hamster ovary
CP450	cytochrome P450
D	thumb subdomain

DOB	5’-(2-phosphoryl-1,4-dioxobutane)
DDR	DNA damage response
DMEM	Dulbecco’s Modified Eagle Medium
DMT	dimethoxytrityl
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DPC	DNA-protein crosslinks
dRP	deoxyribose phosphate
DSB	double strand break
dsDNA	double stranded DNA
DTIC	dacarbazine
dU	2’-deoxyuridine
EMS	ethyl methanesulfonate
FBS	fetal bovine serum
Fe	iron
FEN1	flap endonuclease I
H-bond	hydrogen bond
HKL	honokiol
hNEIL	human Nei-like DNa glycosylase
hNTH1	human endonuclease III protein 1
hOGG1	human 8-oxoguanine DNA glycosylase 1
HO•	hydroxyl radical
HR	homologous recombination

ICL	interstrand crosslinks
IC_{50}	half maximal inhibitory concentration
IR	ionizing radiation
K72	Lys72
KA	koetjapic acid
K_D	dissociation constant
K_I	concentration of the inactivator at half-maximal inactivation
k_{inact}	maximal rate constant for inactivation
k_{SS}	steady state rate constant
k_{ST}	single turnover rate constant
L	2-deoxyribonolactone
LA	linoleic acid
LB	Luria broth
LP-BER	long-patch base excision repair
LPS	lipopolysaccharide
Lys	lysine
Me^3 dA	N^3-methyl deoxyadenosine
MOA	mechanism of action
MEF	mouse embryonic fibroblast
Mg^{2+}	magnesium ion
MGMT	O^6-methylguanine-DNA-methyltransferase
MMR	mismatch repair

MMS	methyl methanesulfonate
MNNG	*N*-methyl-*N*-nitrosoguanidine
MNU	*N*-methyl-*N*-nitrosourea
MS	mass spectrometry
MutYH	human MutY DNA glycosylase
N	finger subdomain
NA	nervonic acid
NBS	*N*-bromosuccinimide
NCP	nucleosome core particle
NCS	neocarzinostatin
NER	nucleotide excision repair
NHEJ	non-homologous end joining
NSD1	nuclear receptor-bidning SET domain protein 1
nt	nucleotide
Nv	nitroveratryl
O^6Me dG	O^6-methyl deoxyguanosine
PA	pamoic acid
PAR	poly-ADP ribose
PARP	poly (ADP ribose) polymerase
PARPi	PARP inhibitors
PNK	polynucleotide kinase
Pol β	DNA polymerase β

Pol δ	DNA polymerase δ
Pol ε	DNA polymerase ε
Pol ι	DNA polymerase ι
Pol λ	DNA polymerase λ
Pol μ	DNA polymerase μ
Pol θ	DNA polymerase θ
ROS	reactive oxygen species
RNAi	RNA interference
RSH	thiol
SMUG1	single-strand selective monofuncation uracil DNA glycosylase
SP-BER	short-patch base excision repair
SSB	single strand break
ssDNA	single strand DNA
TAMRA	carboxytetramethylrhodamine
TBS	tris-buffered saline
TBST	tris-buffered saline with tween 20
TCA	trichloroacetic acid
TDG	thymine DNA glycosylase
TdT	terminal deoxynucleotidyl transferase
TFA	trifluoroacetic acid
Tg	thymine glycol
THF	tetrahydrofuran

TI	therapeutic index
TMZ	temozolomide
TNR	trinucleotide repeat
TRF2	telomeric repeat-binding factor 2
UBER	universal base excision repair
UDG	uracil DNA glycosylase
V(D)J	Variable Diversity Joining
WT	wild type

1. Introduction

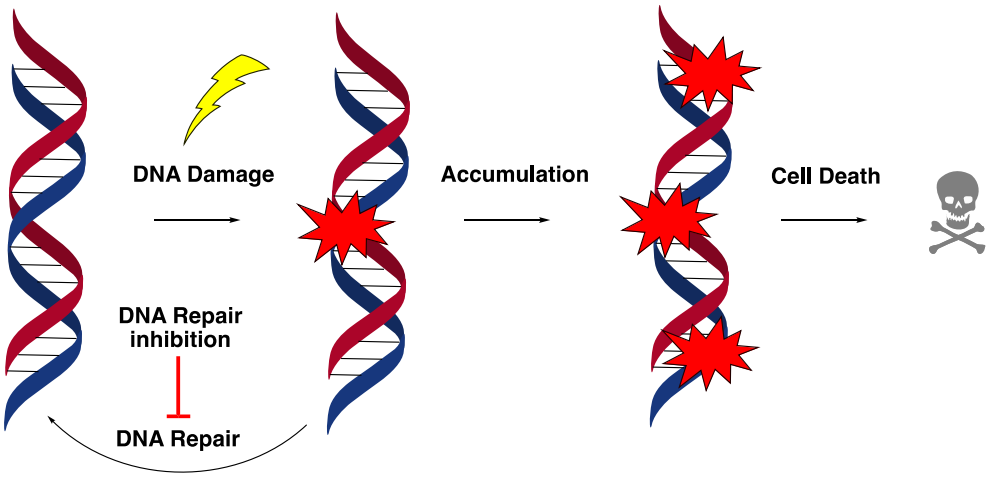
A major approach to kill cells is by attacking nucleic acids using ionizing radiation (IR) or chemotherapeutics, such as bleomycin (BLM), temozolomide (TMZ), or cisplatin. While these cancer drugs have different mechanism of actions (MOAs) to induce cell death, they target deoxyribonucleic acid (DNA). Cancer cells are characterized by unregulated cell proliferation. They no longer possess the normal checks and balances that control and limit cell division. As a result, cancerous cells divide and reproduce much faster than normal cells. Many chemotherapeutic drugs are most effective at killing cells that are rapidly dividing.

Unfortunately, a major challenge of chemotherapeutics is a lack of selectivity. Specifically, many drugs cannot distinguish between cancerous cells and healthy cells. This leads to undesirable non-specific toxicity and as a result, there is a need for the development of selective chemotherapeutics. A drug's therapeutic index (TI) is a quantitative measurement of its relative safety. TI is a comparison of the amount of the drug that causes a therapeutic effect to the amount of drug that induces toxicity. DNA polymerase β (Pol β) is a DNA repair enzyme involved in the repair of small, non-bulky DNA lesions and imparts resistance to drugs that induce DNA lesions. Selective inhibitors of this enzyme that are effective in cells do not exist. The goal of this research is to identify a small molecule inhibitor(s) of Pol β to improve the selectivity of cancer treatment either by enhancing an agent's TI by increasing the cytotoxicity of DNA damaging agents or by inducing synthetic lethality in repair-deficient cancer cells.

Many chemotherapeutics can be classified into two very broad classes: targeted therapies and standard chemotherapy. Targeted cancer therapies block the growth and spread of cancer by interfering with specific molecules or enzymes. These are often chosen deliberately based on extensive knowledge of protein expression levels in individual tumors and validation of the

molecular target. On the other hand, standard chemotherapy takes a broader approach. While this depends on the individual drug, a major target for many standard chemotherapeutics is DNA. Standard chemotherapeutics, such as IR or BLM, induce DNA damage. DNA repair thwarts the effects of DNA damaging agents and reduces the cytotoxicity of drugs, often requiring higher and more frequent doses (Scheme 1). While some cancers exhibit deficiencies in specific repair pathways, others overexpress DNA repair enzymes. An understanding of these changes in cancer cells can guide or hinder approaches involving DNA damage. DNA repair is a crucial mechanism of drug resistance, especially in cells that have higher expression levels of repair proteins than in noncancerous cells.[1] To combat this resistance, DNA repair inhibition is an attractive route to enhance the cytotoxicity of DNA damaging drugs.[1–4]

Scheme 1. DNA repair thwarts the effects of DNA damaging agents.



Maintenance of the genome is vital and regulated by balanced expression levels of DNA repair enzymes. Genomic instability is a characteristic of most cancer cells. Cancer frequently results from damage and mutations in multiple genes controlling cellular division and tumor suppressors.[5] In a sense, higher mutation rates in cancer cells are advantageous in terms of survival and proliferation because genomic instability can give rise to mutations that upregulate genes associated with cancer progression and downregulate genes associated with cancer suppression.[5]

Healthy cells maintain a balanced DNA damage response (DDR) with back-up enzymes and pathways, whereas cancerous cells often exhibit increased or decreased levels of certain repair pathways. Therefore, normal cells can accommodate for a deficiency of one pathway better than cancer cells. Cancerous cells must maintain genomic integrity to avoid apoptosis. Some cancers, especially those that upregulate mutagenic repair pathways or are deficient in a specific repair pathway, rely more heavily on a subset of repair pathways in comparison to normal cells.[6,7] By this logic, DNA repair inhibition is more detrimental in cancer cells than in healthy cells by utilizing synthetic lethality (Section 2.4). Synthetic lethality occurs between two genes when the disruption of both yields cell death, while the loss of one does not. Synthetic lethality is induced by inhibiting an enzyme that is toxic only in cancer cells that are deficient in a different pathway.

DNA repair inhibitors are used clinically in cases in which tumors have higher resistance to DNA damaging agents. For instance, the common alkylating agent, TMZ induces several DNA adducts including, O^6-methyl deoxyguanosine (O^6Me dG) (Figure 1).[8] O^6-alkylguanines are mutagenic because they pair with thymine and result in a GC to AT transition.[9] This lesion is directly removed from DNA by O^6-methylguanine-DNA-methyltransferase (MGMT).[8] MGMT is overexpressed in 80% of brain tumors and is a therapeutic target especially for tumors resistant to alkylating agents.[9] Furthermore, treatment of MGMT inhibitor, disulfiram, confers resistance to alkylating agents in brain tumors.[9]

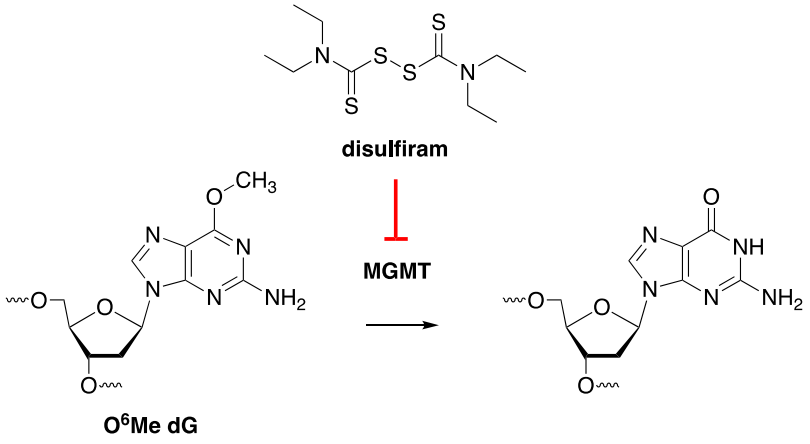


Figure 1. MGMT inhibitor, disulfiram, blocks repair of O^6Me dG.

Pol β has been considered an attractive target for DNA repair inhibition. Pol β is the primary polymerase of an important DNA repair pathway, base excision repair (BER).[10] Pol β is overexpressed in several types of cancer cells, suggesting that the lack of Pol β may be more detrimental to cancer.[11] Pol β is postulated to be synthetic lethal in homologous recombination deficient cells.[12] If confirmed, this would make Pol β a better target for monotherapy in BRCA1/2-deficient cancers (Section 3.6). Multiple mutant forms of Pol β have been identified in a large percentage of tumors, which introduces the potential for identifying inhibitors selective for mutants associated with cancer over WT enzyme (Section 2.5).[13] Enhancing selectivity to cancer cells over healthy cells improves the ratio of therapeutic effects compared to the induced toxicity, in turn augmenting a drug's TI. Significantly, a lack of Pol β activity has been shown to enhance the cytotoxicity of DNA damaging agents.[14–19]

Several inhibitors of Pol β have been identified but lack crucial properties required for clinical use (Section 2.6).[20–28] The current limitations of known Pol β inhibitors include a lack of selectivity for Pol β over other polymerases, lack of potency, and undefined binding region. As a result, there is a need for the development of Pol β inhibitors that have been characterized completely (i.e. kinetics and binding) and are both selective and potent (Section 3.3-3.4).

Previously, it was observed that oxidized abasic lesions, products of oxidative damage, irreversibly inactivate Pol β (Figure 2).[29,30] Based on this finding, small molecules that mimic this form of damage were identified that irreversibly inhibit Pol β.[26,27] In the research presented, the selective binding to the unique domains of Pol β was evaluated (Section 3.3). Pol β was selectively and irreversibly inhibited by a small molecule(s) that mimics oxidized abasic lesions (Section 3.4).

The Pol β inhibitor was synthetic lethal in BRCA1-deficient cells (Section 3.6). An inhibitor library was screened against mutant forms of Pol β but did not yield any selective hit compounds due to a fault in the screening method (Section 3.5). A plan to overcome this flaw was outlined (Section 4). If completed successfully, this would represent the first report of an inhibitor selective for a cancer-specific Pol β variant. This research contributes to the identification of chemical agents that distinguish between cancerous and healthy cells in order to reduce toxicity to healthy cells.

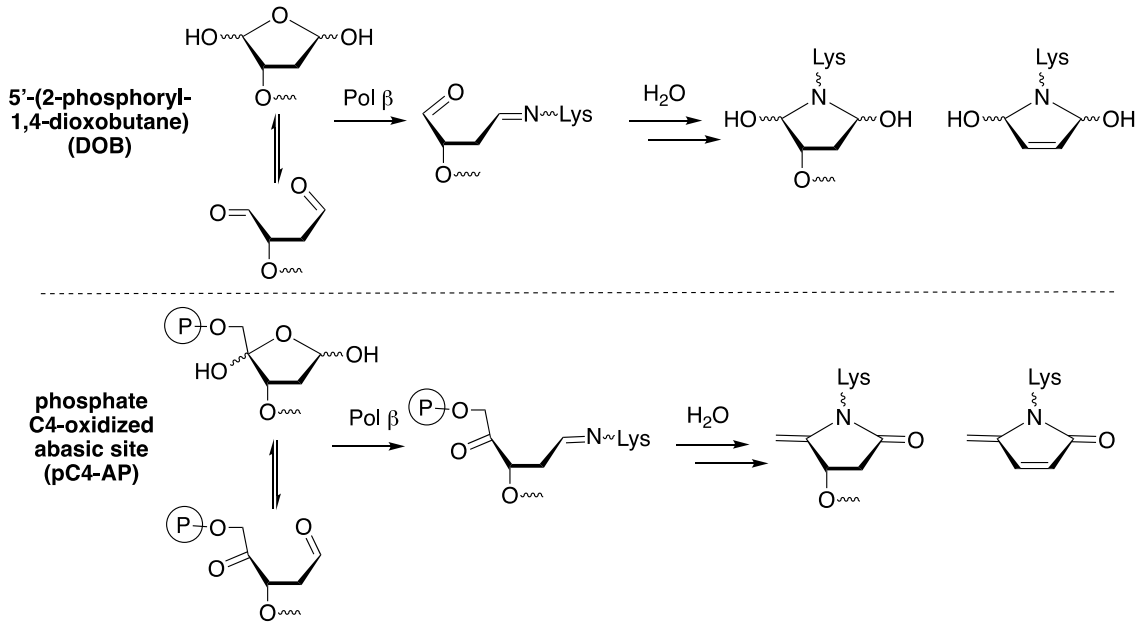


Figure 2. Oxidized abasic lesions that inhibit Pol β.

2. Background

2.1 Cancer Therapeutics

Cancer is the second leading cause of death worldwide. It is a collection of diseases that can affect essentially any part of the body. Cancer is not caused by any single factor but often arises from multiple contributions, including genetic and environmental factors. Due to the complexity of the disease, there are various treatments for cancer, such as surgery, ionizing radiation (IR), and chemotherapeutics. Treatment must balance efficacy and toxicity. A major hurdle for cancer drugs is a lack of selectivity resulting in non-specific toxicity. Therefore, the dosage of cancer drugs is often limited by toxicity and drug resistance conferred by the overexpression of repair enzymes in some cancerous cells.[3,4]

Many chemotherapeutics induce cell death by damaging deoxyribonucleic acid (DNA) (Scheme 1). Several clinically relevant standard chemotherapeutics that generate DNA damage can be characterized into three types: alkylating agents (Section 2.1.1), antimetabolites (Section 2.1.2), and antitumor agents (Section 2.1.3). These agents are often agnostic towards cell type. Specifically, these drugs do not distinguish between tissues or healthy vs. cancer cells and, therefore, are generally associated with high toxicity. It is important to note that this list is not exhaustive, and many selective drugs exist that do not fit into these three categories.

2.1.1 Alkylating agents

Temozolomide (TMZ), dacarbazine (DTIC), and cisplatin (Figure 3) are examples of clinically relevant alkylating agents.[4] Alkylating agents modify DNA in a way that blocks or changes replication and/or transcription.[31] For example, N^3-methyl deoxyadenosine (Me^3 dA) blocks replication while O^6-methyl deoxyguanosine (O^6Me dG) does not block replication but rather preferentially base pairs with thymidine, leading to a mutation (Figure 4).[9,31] The site and

identity of alkylation depends on the alkylating agent. For instance, the products of TMZ and DTIC do not resemble the cross-link formed from cisplatin.

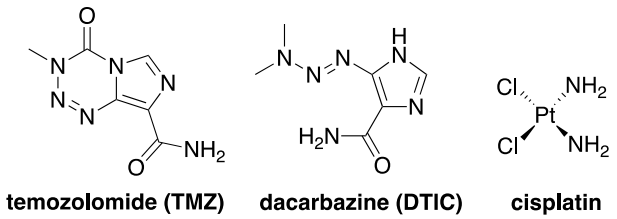


Figure 3. Structures of alkylating agents.

DTIC is oxidized by cytochrome P450 (CP450) in the liver. TMZ and DTIC undergo hydrolysis and decarboxylation to a common intermediate (**I**) (Figure 4). A diazonium ion is released that reacts with a nucleophile in DNA.[8] This results in the methylation of bases, which in some cases lead to the accumulation of apurinic/apyrimidinic (AP) sites in the DNA. TMZ is used to treat brain tumors and gliobastoma and DTIC is often used to treat skin cancer and Hodgkin's lymphoma.

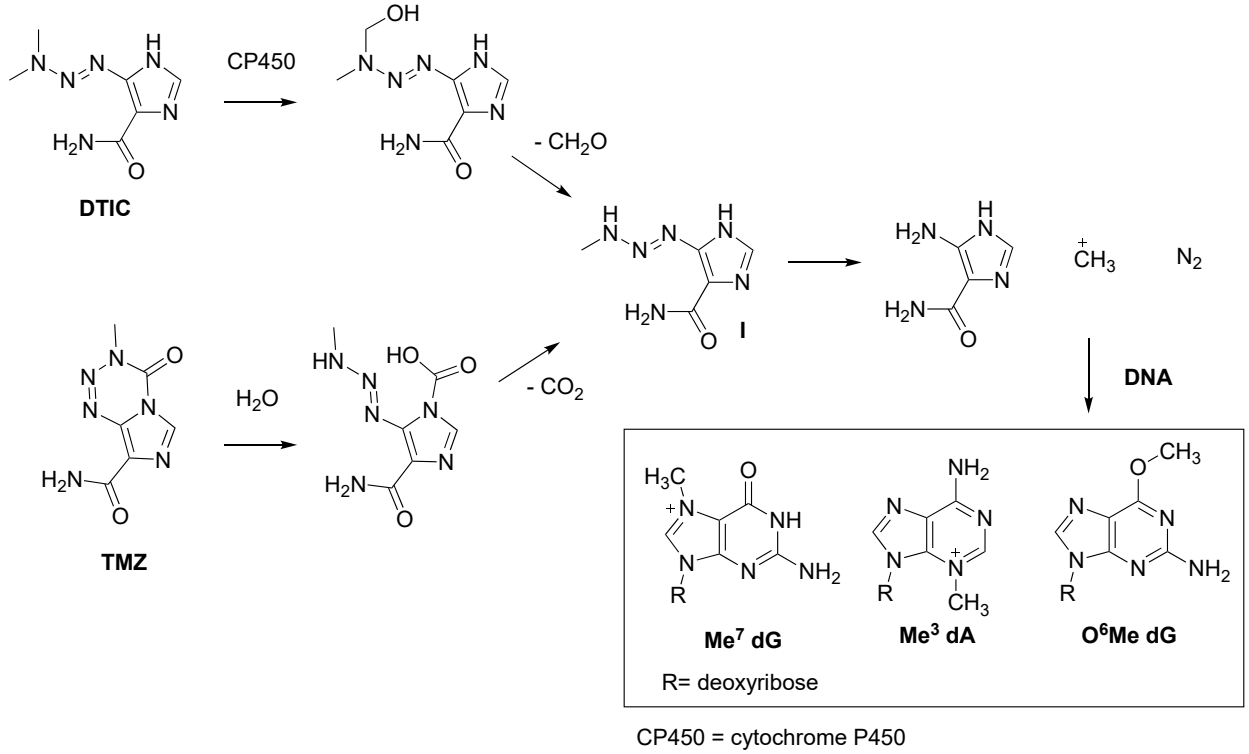


Figure 4. DNA damage by TMZ and DTIC.

Cisplatin is an alkylating-like platinum agent that reacts with deoxyguanosine in DNA. Depending on the local DNA sequence, this forms a cross-link between two adjacent guanines in the same strand (Figure 5). This forms a severe kink in the DNA, preventing normal cellular processes, like replication. Cisplatin therapy is effective in the treatment of testicular, ovarian, bladder, and cervical cancers.[3,32]

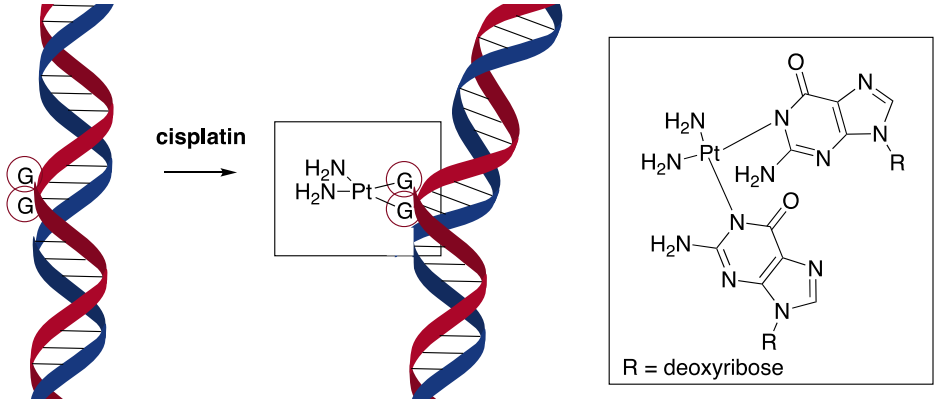


Figure 5. Cisplatin DNA adduct.

2.1.2 Antimetabolites

Antimetabolites mimic the cell's normal building blocks and substrates to disrupt cellular processes.[3] Some examples of antimetabolites include methotrexate and 5-fluorouracil.

Methotrexate inhibits dihydrofolate reductase, an enzyme that converts dihydrofolic acid to tetrahydrofolate (Figure 6). Tetrahydrofolate is a cofactor of many enzymes, including those involved in the biosynthesis of amino acids and nucleotides. Therefore, methotrexate blocks the synthesis of DNA by reducing the production of nucleotides. It is the primary antifolate used in chemotherapy to treat breast, ovarian, bladder, and head and neck cancers.[33]

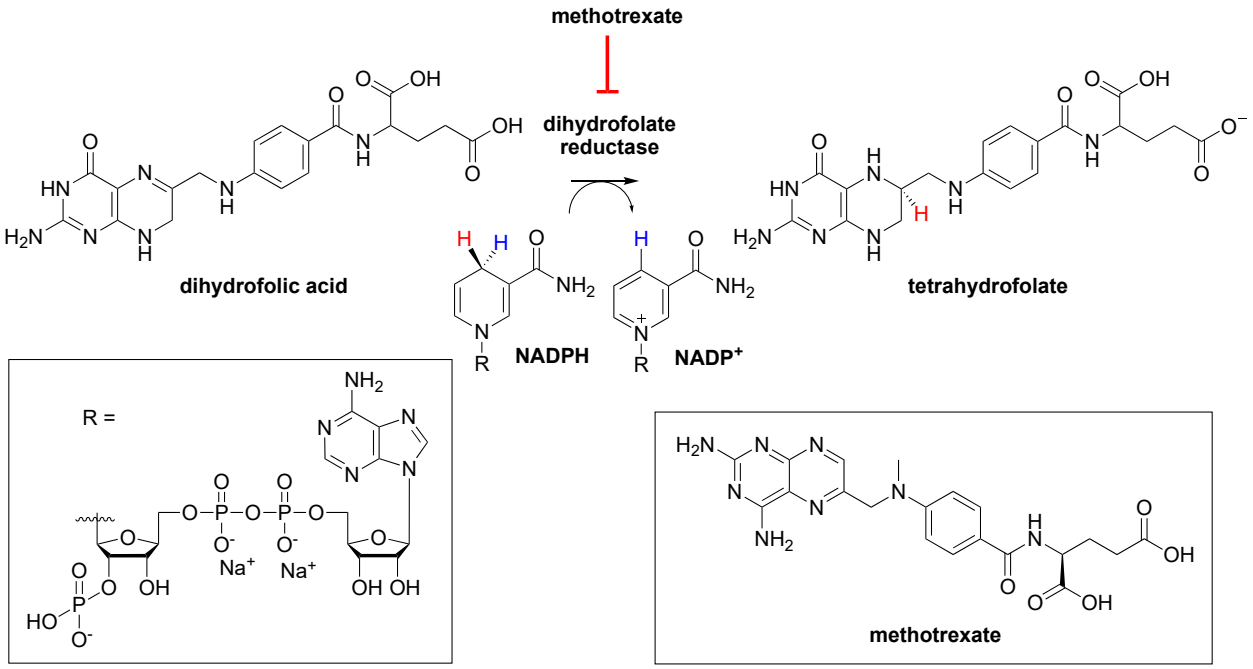


Figure 6. Methotrexate mechanism of action.

The antimetabolite, 5-fluorouracil (5FU), is used to treat breast, bowel, skin, and pancreatic cancer. This agent mimics the structure and size of natural nucleobases, uracil and thymine (Figure 7). Enzymes involved in the biosynthesis of nucleotide triphosphates accept 5FU. Specifically, thymidylate synthase, the enzyme responsible for the synthesis of thymidine is inhibited by 5FU. This interaction blocks the synthesis of thymidine, one of the nucleosides required for DNA replication.[34]

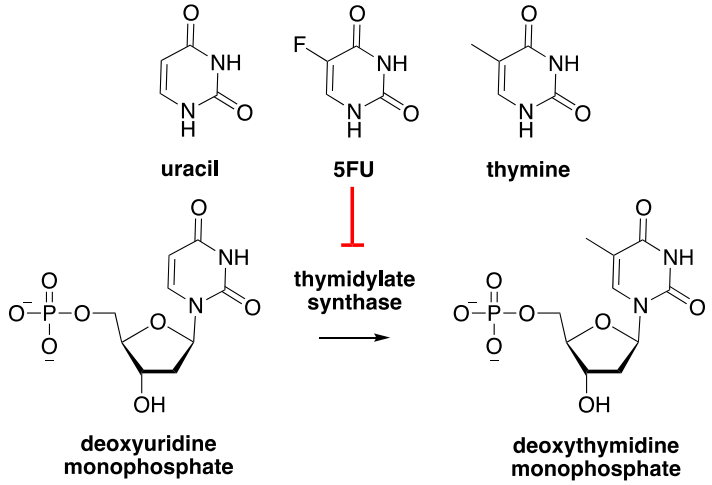


Figure 7. 5FU mimics uracil and thymine.

2.1.3 Anti-tumor antibiotics

Another popular class of chemotherapeutics is anti-tumor antibiotics. Within this class are two major subclasses: anthracyclines and non-anthracyclines. Anthracyclines interfere with enzymes involved in DNA replication by binding to the DNA. One of the most well-known anthracyclines is doxorubicin (Figure 8). Doxorubicin is a highly potent cancer drug and is used to treat a wide range of cancers, including liver, kidney, leukemia, gastric, and several others. Doxorubicin intercalates DNA and inhibits the progression of topoisomerase II, an enzyme responsible for relaxing supercoils in DNA that is required for DNA replication.[3]

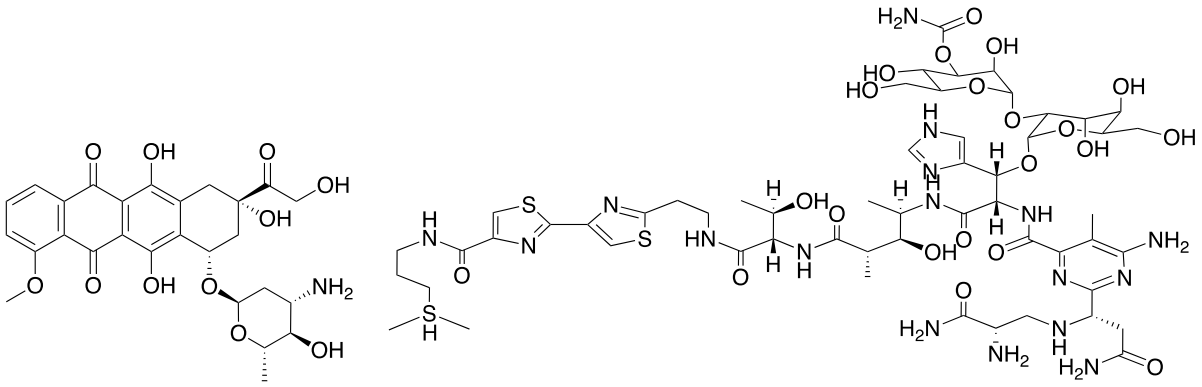
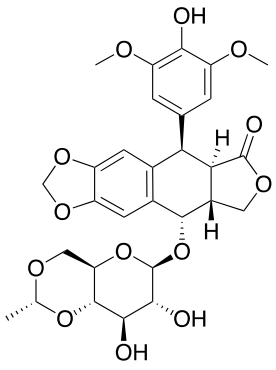


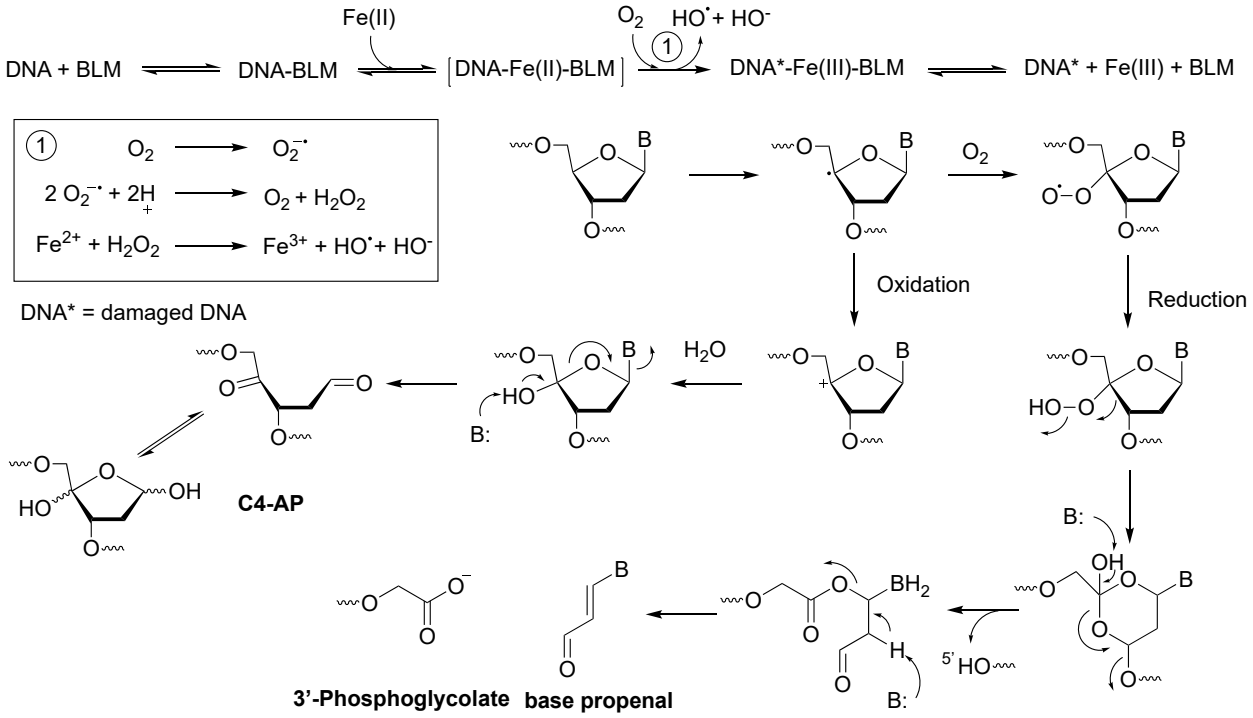
Figure 8. Structures of anti-tumor antibiotics.

Non-anthracyclines damage DNA and induce cell death by some other mechanism. An example of this subclass is bleomycin (BLM) (Figure 8).[35,36] There are a few derivatives in the BLM family (e.g. pingyangmycin, Blenoxane) that oxidatively damage DNA. It is typically used for squamous cell carcinomas, germ cell tumors, Hodgkin's lymphomas, and testes tumors. BLM can be more effective when applied in combination treatment with other chemotherapeutics.[3,37] For example, the cure rate for disseminated testicular cancer increased from 5% to 90% when cells were treated with a combination of BLM, etoposide



(a topoisomerase inhibitor, Section 2.1.4), and cisplatin.[3] BLM forms a reactive radical in a metal-dependent manner.[35,36] The resultant radical abstracts the C4'-hydrogen from the sugar moiety and results in oxidized AP lesion, C4'-oxidized AP lesion (C4-AP), or a single strand break (SSB) (Scheme 2).[36,38]

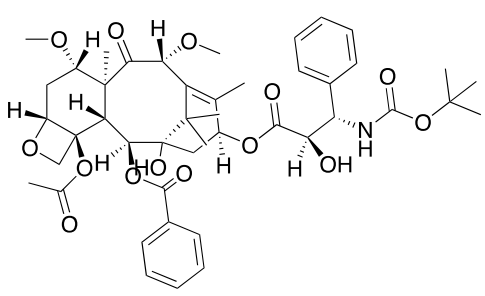
Scheme 2. Mechanism of BLM products.



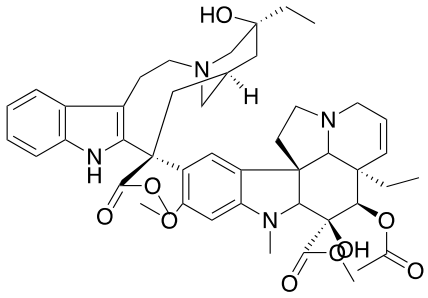
2.1.4 Other classes of chemotherapeutics

Many other classes of chemotherapeutics exist, such as topoisomerase inhibitors, mitotic inhibitors, and several others.[3] Topoisomerases are enzymes that participate in DNA unwinding that is necessary for the initiation of replication.[39] Topoisomerase inhibitors, such as etoposide, are used to treat certain leukemias, as well as lung, ovarian, colorectal, and pancreatic cancers.[3,39] Mitotic inhibitors typically inhibit enzymes that are required for cell division.[40] Examples of mitotic inhibitors include Cabazitaxel (a derivative of taxol) and Vinblastine that are usually used in the treatment of breast, lung, myelomas, lymphomas, and leukemias.[3,40] While there are many

other examples of clinically relevant chemotherapeutics, these drugs typically interrupt normal cellular division and replication, resulting in cell death and reducing proliferation in tumor cells.



Cabazitaxel



Vinblastine

2.2 DNA Damage

DNA damage is an inescapable aspect of life. Every cell's DNA experiences some form of damage at least 10,000 times per day.[41] Multiple forms of DNA damage exist, each with different consequences, toxicity, and mechanisms for repair. Genomic integrity is threatened by endogenous sources and exogenous sources. Endogenous sources include spontaneous hydrolysis of the glycosidic bond and reactive oxygen species (ROS) that our own metabolism generates during cellular respiration. Exogenous physical and chemical sources include IR due to sun exposure and industrial pollutants. The damage can result in modification or loss of the nucleobase, oxidation of the deoxyribose sugar, and/or a break in the phosphodiester backbone (Figure 9).

It is crucial to understand how each lesion is formed, how frequently it occurs, and how it affects cellular processes. While some lesions are repaired efficiently, others lead to more dangerous forms of damage.[42,43]

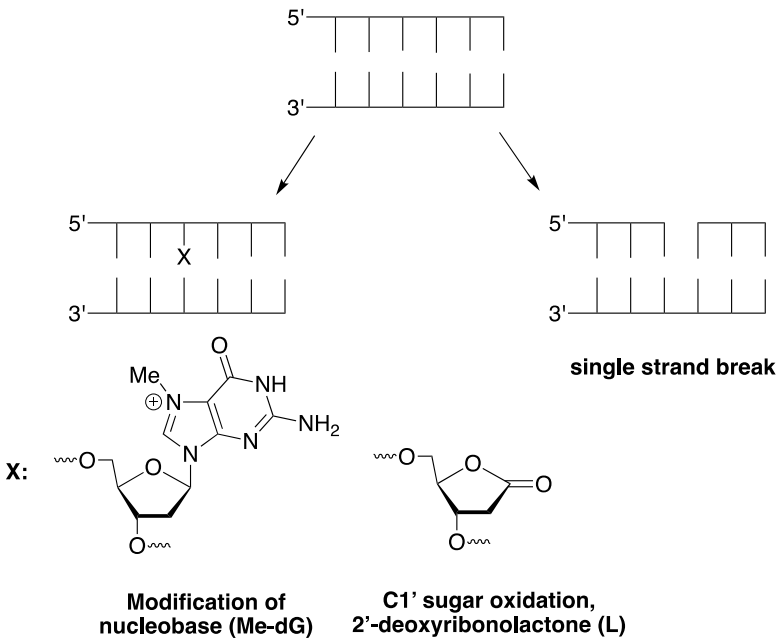
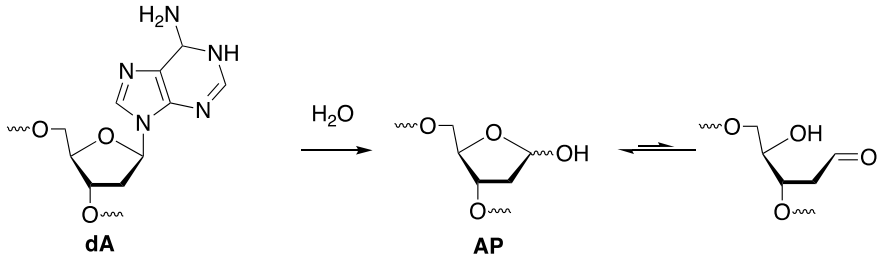


Figure 9. Representative types of DNA damage.

2.2.1 Abasic Sites

AP sites are formed from hydrolysis of the nucleobase moiety of DNA (Scheme 3). The rate of depurination is very slow ($k = 4 \times 10^{-9}$ s^{-1} at 70° C and pH 7.4, $t_{1/2}$ = 730 yrs).[44] Despite the slow rate, 10,000-20,000 AP lesions are formed in a cell every day due to the large size of the human genome, making it the most commonly produced DNA lesion.[44,45] The rate of hydrolysis increases in response to treatment with alkylating agents because alkylated bases are better leaving groups.[4] An AP lesion is the product of N-glycosylases that are involved in the repair of damaged nucleotides (Section 2.3).

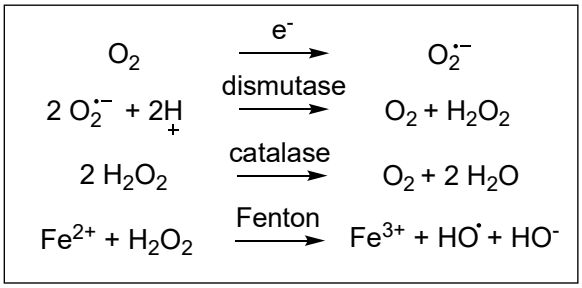
Scheme 3. Formation of AP sites.



AP sites are efficiently repaired by a process known as base excision repair (BER) under normal cellular conditions (See Section 2.3.1). However, if left unrepaired, AP sites can lead to more deleterious forms of damage, such as interstrand crosslinks (ICLs) with the exocyclic amine of dG and dA.[46,47] Unrepaired AP sites can lead to increased levels of DNA strand cleavage. The AP site is a cyclic hemiacetal that exists in equilibrium with the ring-opened aldehyde (Scheme 3), which can undergo β-elimination to yield a single strand break (SSB).[48,49] In addition, AP sites block replicative polymerases during DNA synthesis, resulting in fork collapse if not bypassed by a translesion synthesis polymerase (Section 2.3.1.1.3).[50] An AP lesion does not have a Watson-Crick base-pair and, therefore, does not translate information to polymerases during replication. When a polymerase encounters an AP site, an adenosine is most often incorporated opposite the AP site. Therefore, AP is usually a pro-mutagenic lesion.

2.2.2 Oxidative DNA Damage

DNA is frequently damaged by oxidation. Intracellular oxidizing agents are produced by cellular respiration. ROS are produced endogenously (and exogenously) and their formation is initiated by the formation of superoxide anion from the reduction of molecular oxygen.[51] Superoxide dismutase processes the superoxide anion into hydrogen peroxide and oxygen.[52] Hydrogen peroxide is detoxified by catalase. However, in the presence of iron (Fe) hydrogen peroxide generates the highly reactive hydroxyl radical (HO•) via Fenton chemistry.[53]



Radiation therapy also relies on the formation of HO• in tumors from the direct ionization of water by γ-radiolysis.[54] HO• is highly reactive and unselective. It abstracts hydrogens from the sugar moiety and directly damages the nucleobases in DNA by adding to π-bonds. Nucleobase

modification accounts for as much as 90% of HO• reactivity, while sugar oxidation accounts for less than 10%.[55]

2.2.2.1 Sugar Oxidation

Sugar oxidation occurs from the hydrogen abstraction from the sugar ring. Different oxidized abasic lesions are formed depending on the carbon position (Figure 10). Due to the high reactivity of HO•, hydrogen abstraction is typically governed by accessibility.[56] C5'-hydrogen abstraction accounts for more than 50% of the HO• sugar oxidation reactivity while C4'-hydrogen abstraction accounts for about 20%.[56]

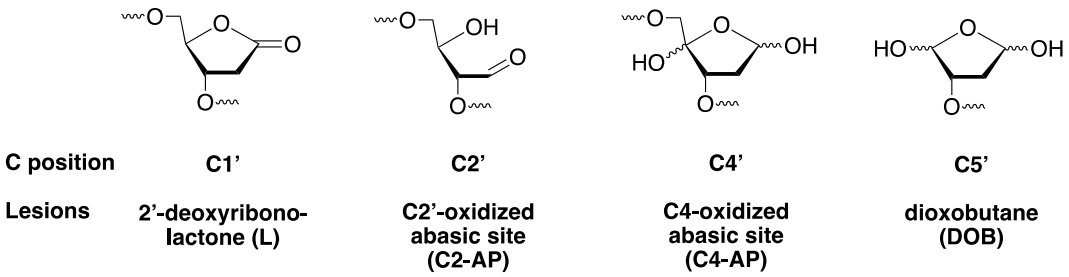
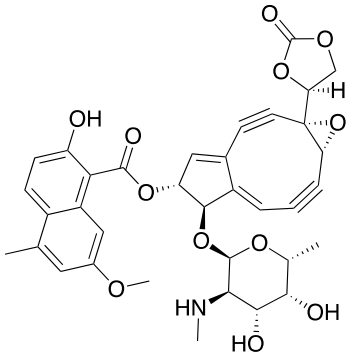
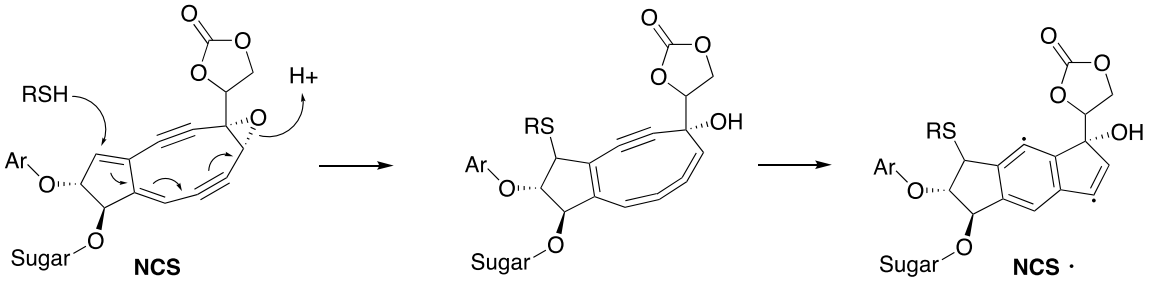


Figure 10. Oxidized abasic lesions.

Sugar oxidation is also caused by exogenous radiomimetics, like BLM and neocarzinostatin (NCS). NCS is a bacterial natural product and is part of the enediyne family. Enediynes feature a macrocyclic ring with a conjugated system containing at least one double and two triple bonds. These minor groove binders are activated by reaction with a thiol (RSH) (Scheme 4). The subsequent cycloaromatization forms a benzenoid diradical (NCS•). This highly reactive intermediate abstracts hydrogen atoms from the sugars of DNA in opposite strands.[37,57]

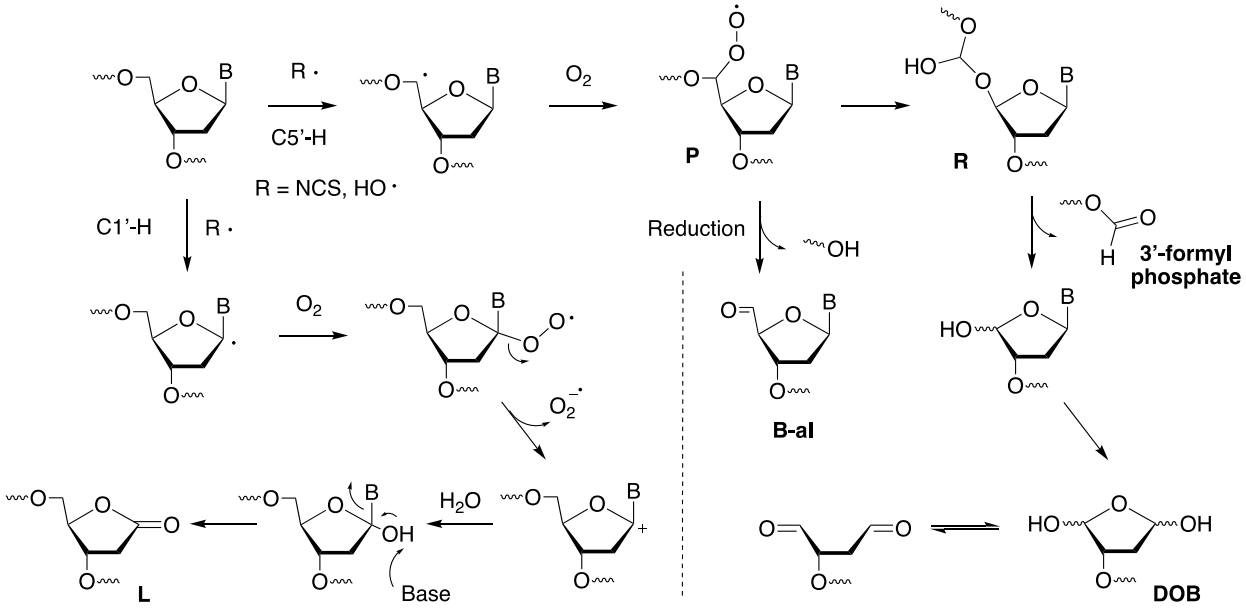


Scheme 4. Activation of NCS.

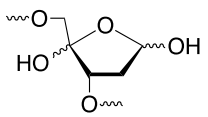


NCS abstracts the C5'-, C4'-, and/or C1'-hydrogen from the 2'-deoxyribose. C1'-hydrogen abstraction yields the oxidized abasic lesion, 2-deoxyribonolactone (L) (Scheme 5, left). The mechanism of the C5'-hydrogen atom abstraction product formation is unclear but was proposed to go through a peroxyl radical intermediate (**P**) (Scheme 5, right).[58] Intermediate **P** primarily leads to the 5'-aldehydes (B-al) (Scheme 5).[58] A minor reaction pathway of **P** involves Criegée rearrangement (**R**) to release the upstream DNA with a 3'-formyl phosphate termini and an oxidized abasic lesion, 5'-(2-phosphoryl-1,4-dioxobutane) (DOB).

Scheme 5. Proposed mechanism of C1'- and C5'- hydrogen atom abstraction product formation.



The other oxidized abasic lesion, C4-AP, is formed from C4'-hydrogen atom abstraction, (Scheme 2). These lesions (i.e. DOB, C4-AP, and L) are produced from NCS treatment. C4-AP produced is more frequently by BLM, where the C4'-position of the sugar is the primary target of BLM.



C4-AP

Oxidized AP lesions can lead to secondary forms of damage that occur due to their spontaneous reactivity. Oxidized AP lesions lead to the formation of more detrimental forms of damage, including ICLs, DNA-protein crosslinks (DPCs), and SSBs.[42,59,60] The acyclic 1,4-dialdehyde of DOB yields ICLs selectively with the exocyclic amine of dA opposite a 3'-adjacent dT.[61] DOB also undergoes spontaneous β-elimination to produce but-2-ene-1,4-dial, which forms pro-mutagenic exocyclic adducts with dA, dC, and dG.[62–64] DOB and pC4-AP form DPCs with and inactivate the DNA repair enzyme, DNA polymerase β (Pol β) (Section 2.3.1.2).[59] In a similar fashion, DOB and C4-AP also covalently modify lysine residues of histone proteins in the context of the nucleosome core particle (NCP).[65,66] In addition, C4-AP forms ICLs with dC in DNA in response to BLM or IR, which was converted to a double strand break (DSB) when incorrectly repaired *in vitro.*[67,68] The biological relevance of these reactions is unclear because the rate of ICL formation with oxidized AP sites is slow *in vitro.* Despite the low probability of ICL formation, a low frequency of ICL events can still present a negative biological effect. The detection of ICL formation between C4-AP and dC in cells support it's biological relevance.[67]

2.2.3 Single strand breaks

The formation of SSBs in DNA is critically important in understanding DNA damage. SSBs are often formed from the processing and inherent reactivity of less deleterious damage (e.g. modified bases, sugar oxidation).[43] For example, the major product of C4' sugar oxidation is a single strand break with an unnatural terminus (i.e. 3'-phosphoglycolate). If left unrepaired, SSBs

can lead to a more dangerous DSB upon DNA replication. Therefore, proper repair of SSB is essential for maintaining genomic integrity.

2.3 DNA Repair

To combat the toxicity of multiple forms of DNA damage, several DNA repair processes exist that focus on the removal and repair of limited sets of modifications (Figure 11). BER repairs non-bulky modifications to the nucleobase that may arise from oxidation, alkylation, or deamination, as well as AP and, albeit less efficiently, oxidized AP lesions.[30,69] Nucleotide excision repair (NER) repairs bulky DNA adducts, ICLs, and nuclebase photodimers via the excision of several nucleotides in the damaged strand of DNA, followed by gap-filling and ligation. Mismatch repair (MMR) recognizes errors of DNA replication that result in a mismatch and excises the mis-incorporated nucleotide and inserts the correct base to restore the DNA. Homologous recombination (HR) and non-homologous end joining (NHEJ) repair DSBs and some types of ICLs.[69,70]

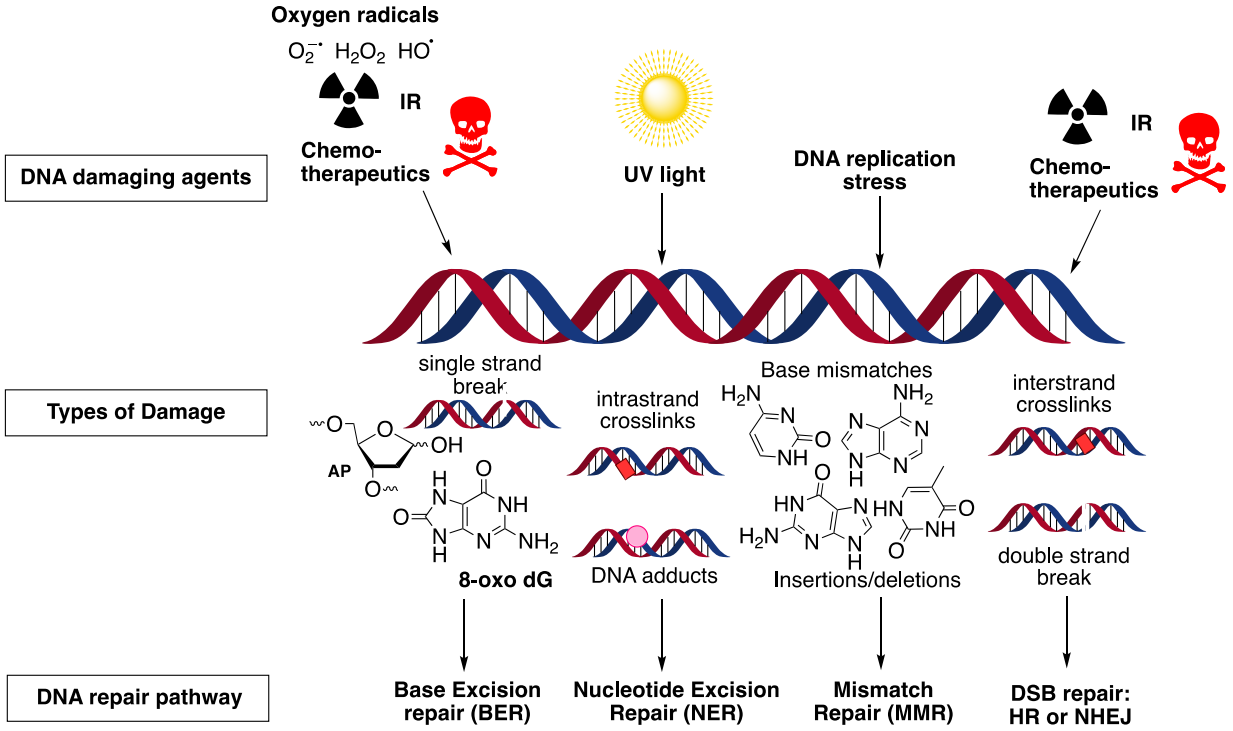


Figure 11. Overview of DNA damage and repair pathways.

2.3.1 Base Excision Repair

There is a misconception that commonly formed lesions, such as small non-bulky adducts, AP sites, and SSBs, exhibit less severe biological consequences in the hierarchy of DNA damage.[43] However, the frequency and abundancy of these lesions present complications for the cell. BER acts on these lesions, is active in all aspects of life, and its contributing enzymes are constituitively expressed in mammalian cells.[71] BER primarily acts on lesions that arise from alkylation and oxidation (Figure 12).[71–75] This includes small methylation adducts, like Me^3 dA, a cytotoxic replication-blocking lesion (Figure 4).[71,74] This also includes damage formed from oxidizing agents, like 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxodG) or thymine glycol (Tg).[76,77] Another BER substrate is deoxyuridine (dU), which results from spontaneous deamination of cytosine or misincorporation during DNA replication.[71,72,74]

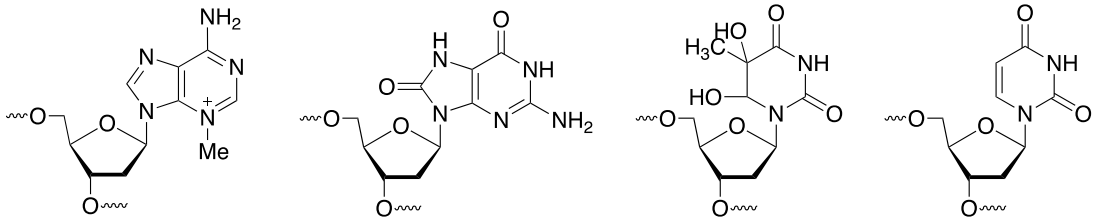


Figure 12. Representative lesions that are repaired by BER.

BER is complex but can be divided into two major mechanisms: short-patch BER (SP-BER) and long-patch BER (LP-BER) (Figure 13). The major difference between the two routes is the number of nucleotides incorporated during the polymerase step. LP-BER incorporates patches greater than one nucleotide (nt) and SP-BER incorporates a single nt.[20,70]

The general mechanism of SP-BER utilizes four major enzymes to carry out five steps (Figure 13a).[70,78] Briefly, an N-glycosylase removes a damaged base by cleaving the glycosidic bond to reveal an AP site. In the case of spontaneous depurination of DNA or oxidized AP site

repair, this first step is unnecessary. The AP site is recognized by AP endonuclease 1 (Ape1). Ape1 cleaves the phosphodiester bond 5' to AP, generating a strand break with 3'-OH and 5'-deoxyribose phosphate (5'-dRP) termini.[79] The following two steps are carried out by Pol β. The Pol β lyase domain removes the 5'-dRP group via β-elimination to yield a 1-nt gap in DNA, which is filled in by Pol β's polymerase domain.[80–82] The resulting nick in the DNA is sealed by DNA ligase.

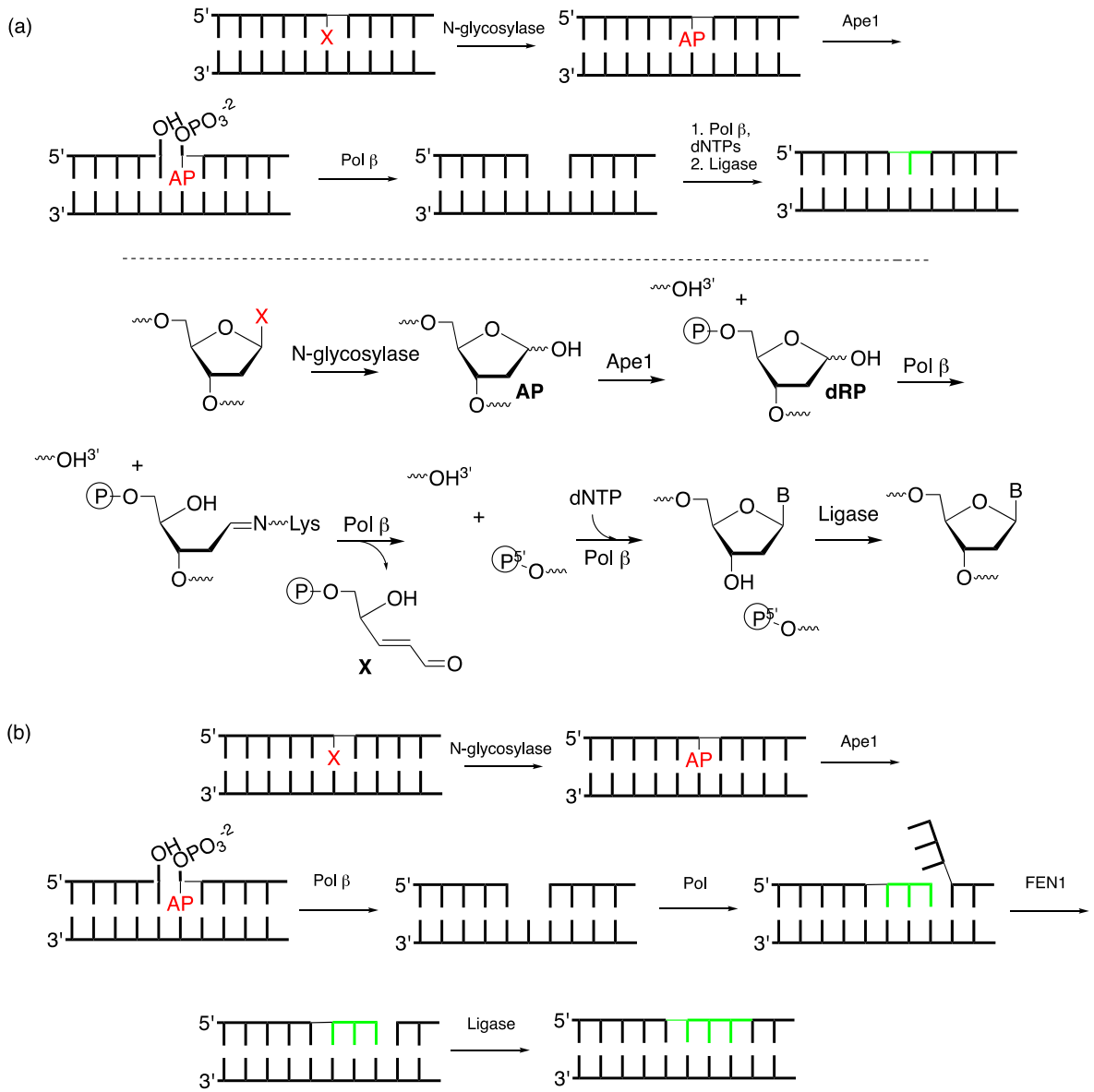


Figure 13. Mechanisms of (a) SP-BER and (b) LP-BER.

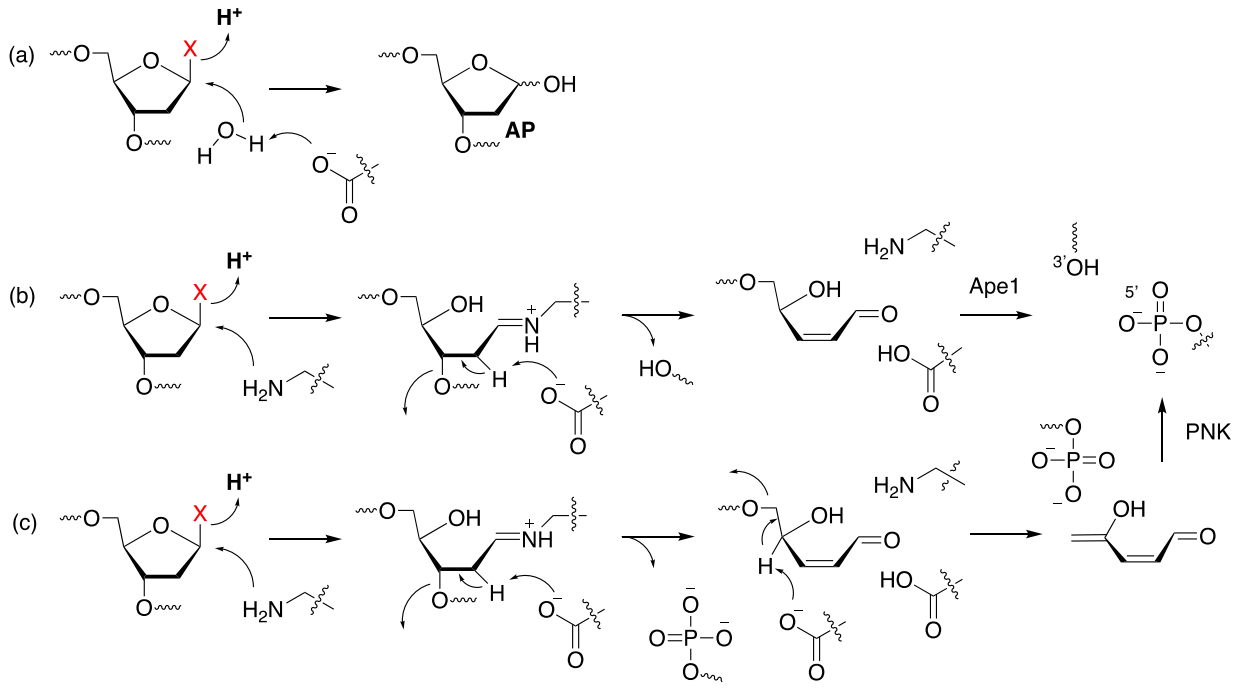
In LP-BER, after the removal of the 5'-dRP termini, nucleotides are incorporated by Pol β (Scheme 13b). It has also been proposed that a higher fidelity polymerase, like Pol δ or ε, may displace Pol β and incorporate multiple nucleotides to create a ~50-nt flap.[20,69] This flap is removed by flap endonuclease 1 (FEN1).[20]

There are at least 11 glycosylase enzymes in mammanlian cells that remove damaged or incorrect nucleobases.[72,83] Glycosylases utilize a base flipping mechanism to excise the damaged nucleobase.[75,83] Most glycosylases utilize base-flipping to search for and replace a damaged base in the genome.[75,84] For example, uracil DNA glycosylase (UDG) inserts the hydrophobic side chain of a leucine residue into the DNA to disrupt the duplex and flip uracil into the active site of the glycosylase.[75] In contrast, thymine DNA glycosylase (TDG) introduces an arginine residue into the minor groove of the DNA to search for a T-G mismatch and upon recognition, flips thymidine out of the duplex for cleavage.[84]

SP-BER can proceed through a different mechanism depending on the identity of the glycosylase. Glycosylases are further classified as monofunctional or bifunctional enzymes. A monofunctional DNA glycosylase (e.g. UDG, SMUG1, TDG, MutYH) catalyzes the hydrolysis of the glycosidic bond (Scheme 6a). Bifunctional glycosylases (e.g. hNTH1, hOGG1, hNEIL) possess AP lyase activity, which would negate the need for some enzymes in the BER pathway (i.e. Ape1 in some instances, Pol β).[82] However, the AP lyase activity of some bifunctional glycosylases (i.e. hNTH1, hOGG1, hNEIL3) is relatively weak and can be overridden by Ape1, reverting to a mechanism of a monofunctional glycosylase (Scheme 6a).[85–88] In the case of a bifunctional glycosylase containing AP lyase activity, other enzymes are required for processing the ends of the strand break to create a suitable substrate for Pol β's polymerase activity. For example, human oxoguanine glycosylase (hOGG1) recognizes 8-oxodG and cleaves the N-

glycosidic bond. The lyase activity of hOGG1 catalyzes a β-elimination, yielding an 3'-dRP and the 5'-phosphate (Scheme 6b). Ape1 hydrolyzes the phosphate backbone to reveal a 1 nt gap with 3'-OH and 5'-phosphate groups. On the other hand, glycosylase activity of human Nei-like DNA glycosylase (hNEIL) proceeds through β-elimination, followed by δ-elimination to leave a 3'-phosphate and 5'-phosphate (Scheme 6c). The 3'-phosphate is processed by the phosphatase activity of polynucleotide kinase (PNK).[82]

Scheme 6. General mechanism of (a) monofunctional glycosylase, (b) bifunctional glycosylase that catalyzes β-elimination, (c) bifunctional glycosylase that catalyzes β,δ-elimination.



In the case of a monofunctional glycosylase, Ape1 hydrolyzes the DNA on the 5' side of an AP site. Ape1 is very efficient at this reaction (single turnover rate constant (k_{ST}) = 850 s^{-1}, steady state rate constant (k_{SS}) = 2 s^{-1}).[89] Ape1 also recognizes oxidized AP sites, albeit at a much lower efficiency (i.e. 5- to 10-fold less relative to AP).[90,91] Ape1 is important for cell viability.

Knockout of Ape1 is embryonically lethal in mice.[92] Furthermore, Ape1 is overexpressed in many cancer cell lines and its inhibition enhances the cytotoxicity of DNA damaging agents.[2,93]

The major BER polymerases in humans are Pol β and DNA polymerase λ (Pol λ).[69] These polymerases belong to the X-family. Other polymerases that possess dRP lyase activity, such as polymerase ι (Pol ι) and polymerase θ (Pol θ), have been implicated in BER.[94] Pol ι knockdown cells have been linked to a deficiency in BER.[95] Separately, Pol θ knockout cells exhibited increased sensitivity to oxidative stress caused by H_2O_2.[96]

BER polymerases must catalyze two of the five steps in BER, with the exception of bifunctional glycosylases that catalyze β,γ-lyase activity. Pol β and Pol λ each possess an 8 kDa lyase domain separate from their polymerase domain. Pol β is a 39 kDa enzyme. Pol λ is larger (68 kDa) and possesses a BRCA1 C Terminus (BRCT) domain, which is utilized for protein-protein interactions in the enzyme's role in NHEJ.[97] Despite these differences, the polymerases share substantial sequence homology (~32%) and the polymerase domains of each enzyme share more than 50% sequence identity.[98–100] In contrast to many other polymerases, Pol β and λ lack a 3'-5' exonuclease domain and have high error rates (10^{-4} compared to 10^{-6}-10^{-8} for replicative polymerases) in mammalian cells.[101] The primary role of these polymerases is not in replication but rather in maintaining genomic integrity.[82]

2.3.1.1 DNA Polymerase β

There are 17 eukaryotic polymerases and 13 of these have roles in DNA repair. Pol β is one of four polymerases in the X-family: Pol β, Pol λ, DNA polymerase μ (Pol μ), and terminal deoxynucleotidyl transferase (TdT).[69] Pol μ and TdT are primarily used in DSB repair, including NHEJ and Variable Diversity Joining (V(D)J) recombination. Pol β is the main polymerase in BER, while Pol λ is considered a back-up polymerase to Pol β and has its own role in other repair

pathways, including NHEJ and V(D)J recombination.[97,102] Pol β consists of two major domains: an 8 kDa amino terminal lyase domain (residues 1-87) and a 31 kDa carboxy terminal polymerase domain (residues 88-334) that are linked together through a protease sensitive region between residues 82-86.[80,82,103]

Binding studies of the two domains of Pol β show that the 8 kDa N-terminal domain has a higher affinity for single strand DNA (ssDNA), while the 31 kDa C-terminal domain binds double stranded DNA (dsDNA), deoxynucleotide triphosphates (dNTPs), and possesses polymerase catalytic activity but no ssDNA binding capacity (Figure 14).[104] The crystal structure of Pol β shows that the 31 kDa domain has three subdomains resembling a right hand: thumb (D), palm (C), and finger (N) subdomains, which has been observed in other polymerases (Figure 14).[105] Pol β has a higher affinity for duplex DNA with nicks bearing 3'-OH/5'-PO_4 termini.[81] On nicked DNA bearing a 5'-dRP group, Pol β performs both the dRP lyase and gap-filling polymerase activity (Figure 13).[10]

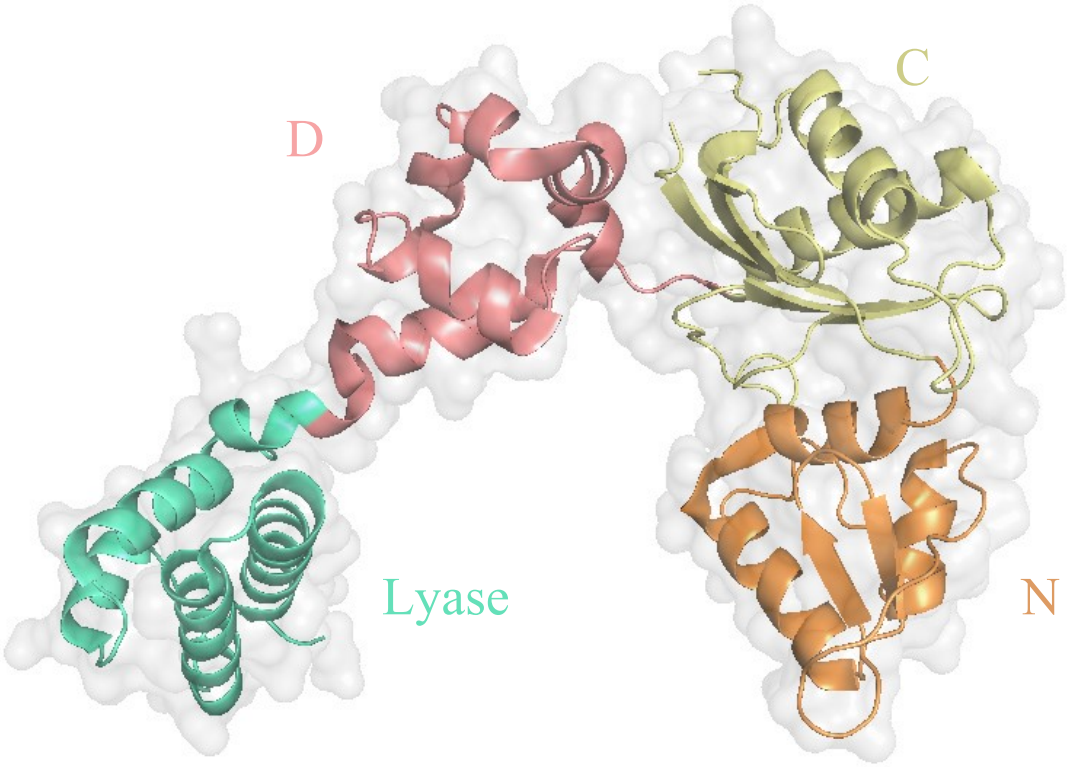
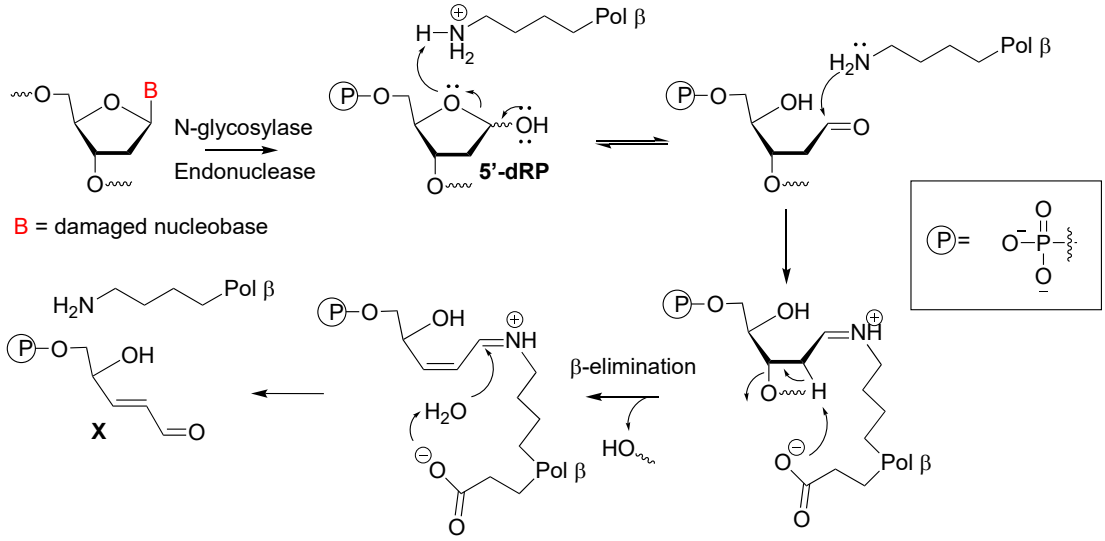


Figure 14. Structure of Pol β (PDB: 1BPD). Key: Lyase = 8 kDa Lyase domain; D = thumb, dsDNA binding subdomain; C = palm, catalytic subdomain; N = finger, dNTP binding subdomain of the 31 kDa polymerase domain.

2.3.1.1.1 Pol β 8 kDa Lyase domain

The 8 kDa domain of Pol β is responsible for releasing the 5'-dRP group via a β-elimination mechanism (Scheme 7). Based on the detection of the elimination product (**X**), researchers proposed that a Schiff base intermediate would form between the C1' position of the AP and the active site lysine (Lys) in the 8 kDa lyase domain.[105] Treatment with $NaBH_4$ trapped the Schiff base intermediate and stabilized the Pol β-DNA DPC.[103]

Scheme 7. Proposed mechanism of 5'-dRP lyase activity.



The AP site is a cyclic hemiacetal that exists in equilibrium with the ring-opened aldehyde. However, the equilibrium lies more heavily on the ring-closed conformation, with only about 1% of the aldehyde present.[48] Schiff base formation requires opening of the dRP ring to the C1' aldehyde. The ring opened form occurs intrinsically during the interconversion between the α- and β-sugar configurations. It is possible that the active site Lys activates the hemiacetal mixture by protonation of the O4' to promote ring opening and formation of the more nucleophilic, neutral Lys.[105] When the iminium ion forms, the pKa of a C2' hydrogen is reduced and deprotonation triggers β-elimination, resulting in the 3'-C-O bond cleavage. Two models were proposed to explain Pol β recognition of the AP site. The first model involves a base-checking mechanism where a nucleophilic Lys residue in the lyase domain scans the DNA by forming hydrogen bonds (H-bonds) with the nucleobases and recognizes the AP site by the lack of H-bonding.[106] The second model proposed that a histidine residue swings into the void created from the missing base in the DNA, which positions the dRP moiety in close proximity to the active site Lys.[105]

Schiff base formation by a Lys residue occurs in other enzymes that possess lyase activity. For example, Lys120 of the bifunctional glycosylase *E.coli* Nth catalyzes β-elimination via schiff base formation. Site-directed mutagenesis experiments suggest that Pol β Lys72 (K72) is the primary amine that forms the Schiff base with the AP site in DNA.[105] The K72R mutant still performs DNA synthesis but is unable to repair dRP sites. Mutants of the 8 kDa domain (K72R and K72Q) completely lost the ability to excise 5'-dRP, even though they were still able to form imine intermediates, albeit to a lesser extent than the wild type (WT) 8 kDa domain.[105,107] In a separate experiment, the K72A mutant only reduced AP repair to 30% that of the WT Pol β. The partially retained lyase activity of K72A mutant indicates that a different Lys residue in the active site may be able to substitute for K72 in its absence to carry out β-elimination.[103] The 8 kDa domain mutant containing K84A did not impair dRP lyase activity but the crystal structure of the enzyme suggests that K84 is the only Lys residue in the active site pocket close enough to the C1' position capable of substituting for K72 in its absence.[108]

Other lyase enzymes also typically use carboxylate or histidine side chains to deprotonate the C2'-H during β-elimination but it is unclear whether Pol β has an appropriate residue that is suitably positioned.[109] Analysis of a model with Pol β in complex with DNA containing a tetrahydrofuran (THF) analog of an AP site suggests Glu26 or Glu71 may be in close proximity to activate a water molecule for C2'-H deprotonation.[107,108]

2.3.1.1.2 Pol β 31 kDa Polymerase domain

The 31 kDa domain is responsible for polymerase activity and depends on the presence of two magnesium ions (Mg^{2+}) that stabilize the transition state and orient the incoming dNTP in the correct position.[110,111] The magnesium ions are observed in the crystal structure bound to all three phosphates of the incoming dNTP, as well as the side chain oxygens of the catalytic triad (Asp190,

Asp192, and Asp256), which are are required for catalysis. Catalysis occurs when the lyase domain closes in relation to the thumb domain, which may happen via favorable hydrophobic and van der Waal contacts between the thumb domain and the sugar moiety of the incoming nucleotide when the correct dNTP is present in the active site of the enzyme (Figure 15).[112]

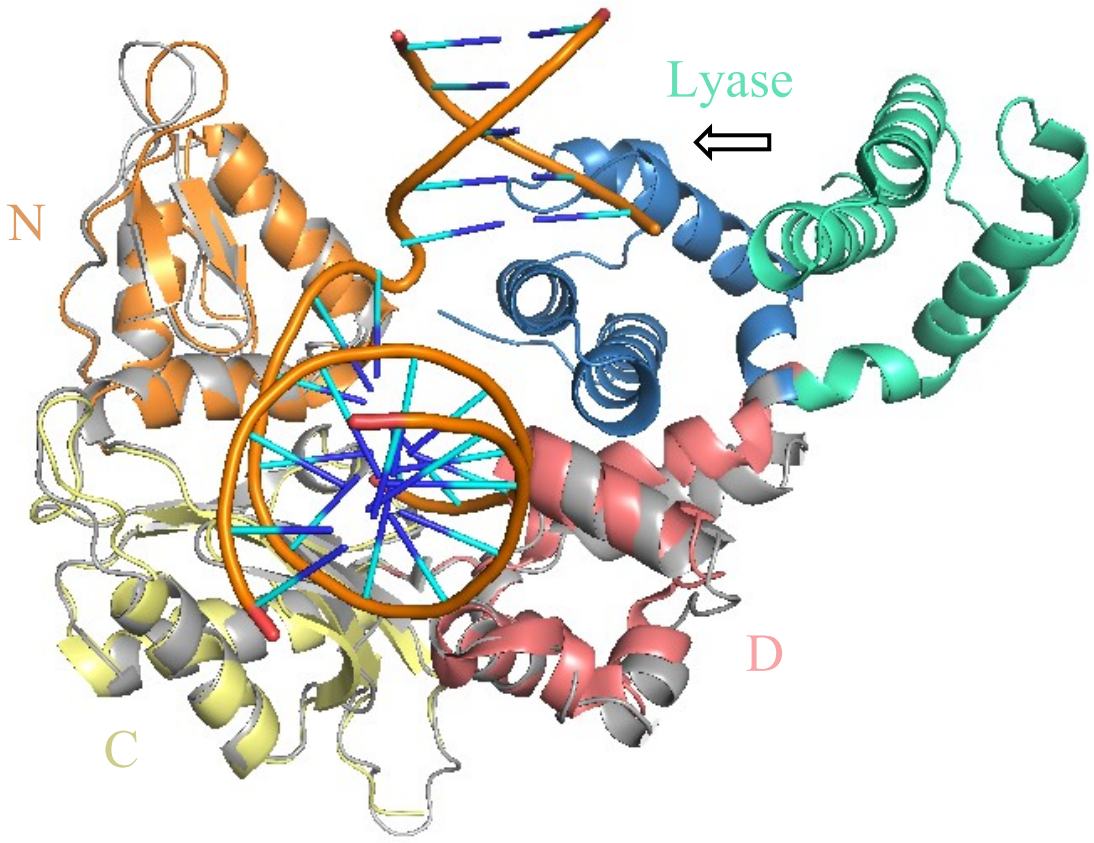
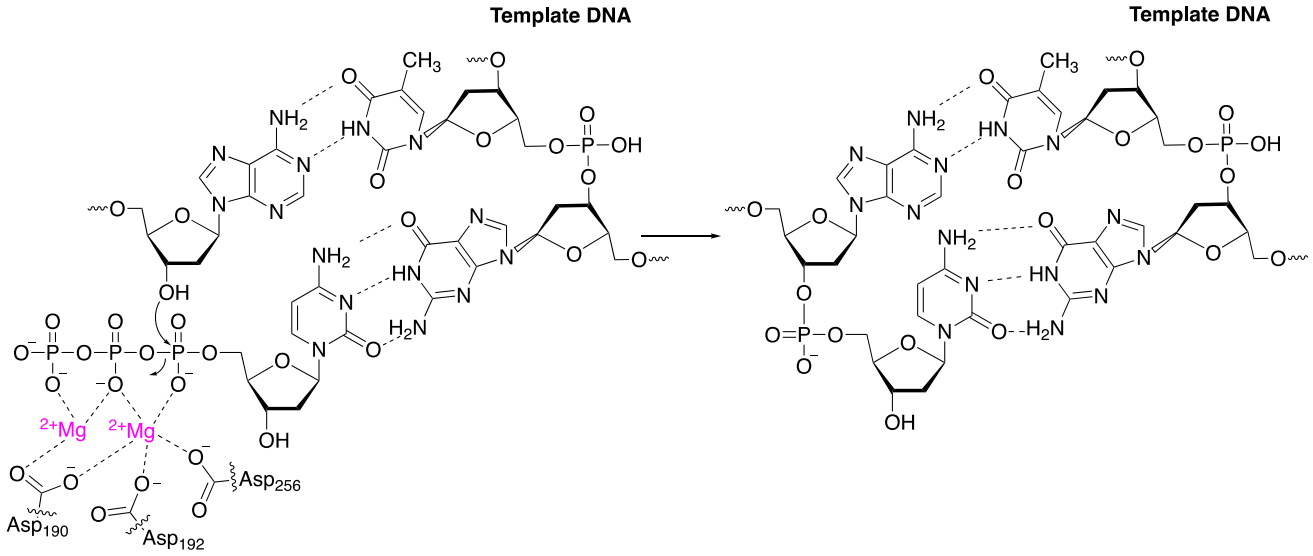


Figure 15. Overlay of DNA-unbound (PDB: 1BPD, teal/pink/yellow/orange) and DNA-bound (PDB: 1BPX, blue/gray) crystal structures of Pol β. Key: Lyase = 8 kDa Lyase domain; D = thumb, dsDNA binding subdomain; C = palm, catalytic subdomain; N = finger, dNTP binding subdomain of the 31 kDa polymerase domain.

When the correct dNTP is inserted into the active site, the 3'-OH is activated and attacks the α-phosphate of the dNTP, elongating the primer strand and a pyrophosphate is released (Scheme 8). Pol β exhibits lower fideltiy compared to replicative polymerases and, therefore, is more likely to make a mistake. In addition to the lack of proofreading activity, the lower selectivity

may be due to the higher acceptance of metal ion cofactors other than Mg^{2+}.[112,113] This was observed in Pol β crystals when non-templated nucleotidyl transfer occurred in the presence of Mn^{2+}.[114] The acceptance of Mn^{2+} by Pol β resulted in greater mutagenicity by promoting greater reactivity. Inclusion of other metal ion cofactors potentially alters the position of the nucleotide in the active site and lead to lower selectivity in distinguishing between nucleotides.[112]

Scheme 8. Mechanism of polymerase activity.



Based on the crystal structure of Pol β complexed with DNA, several residues were identified that are important for binding interactions and catalytic function.[115,116] Lys234, Arg254, Lys280, Arg283, Asn294, Glu295, and Tyr296 interact with the template DNA (Figure 16). Site-directed mutagenesis experiments indicate Asn294 and Glu295 contribute significantly to the polymerase activity of Pol β.[117,118] Of the DNA binding residues, Lys234, and Arg254 are invariant across X-family polymerases.[117] Residues Arg149, Ser180, Arg183, Ser188 interact with the incoming dNTP. Arg183 and Ser180 make direct contacts with the β- and γ-phosphates of the dNTP and, indeed, alanine mutants of each position display a decrease in the catalytic function, suggesting they may play a role in stabilizing the transition state of the reaction.[118] As expected,

mutation of any of the aspartate residues in the catalytic triad (Asp190, Asp192, and Asp256) result in a loss of catalytic activity.[118]

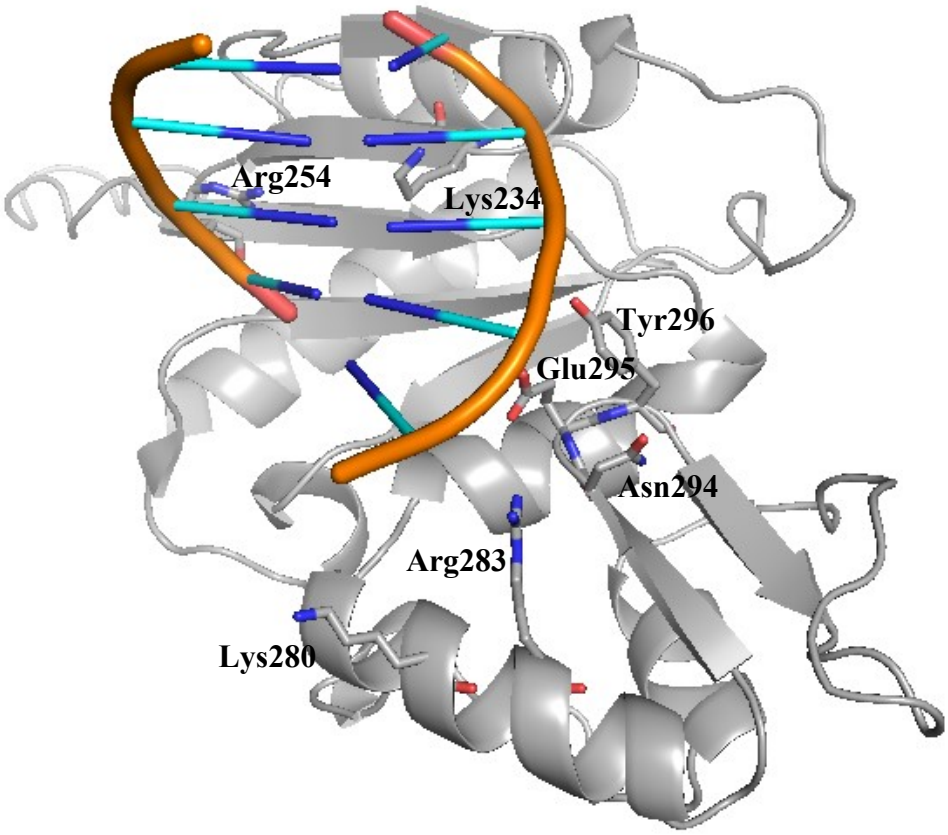


Figure 16. Orientation of DNA binding residues (sticks).

2.3.1.1.3 The effects of Pol β levels in cells

Pol β is crucial for cell viability and its regulation is important for normal cellular functions. In cells that overexpress Pol β, sequestration of telomeric repeat-binding factor 2 (TRF2) leads to premature telomere end fusion.[101] Briefly, telomere end fusion occurs when two chromosome ends are linked together after telomeres lose their protective proteins (i.e. TRF2) and leads to problems with proper segregation during mitosis and meiosis (Figure 17).[119] An overabundance of Pol β could recruit TRF2 away from its normal cellular function and lead to chromosomal instability.[101]

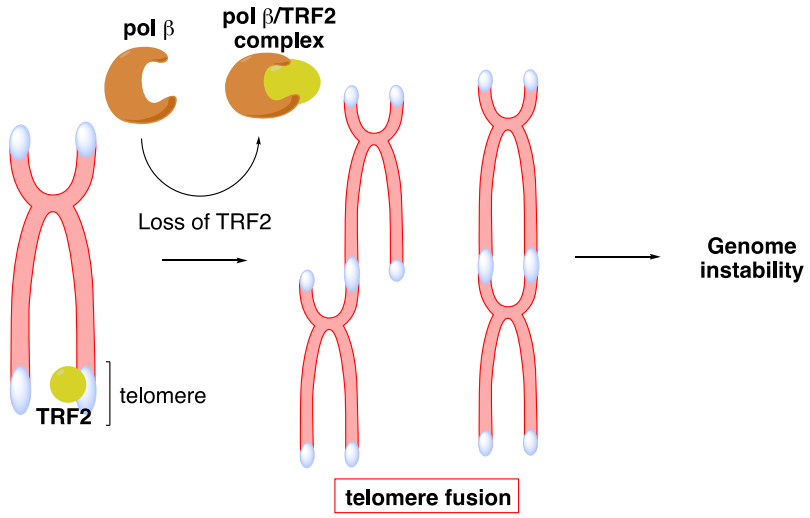


Figure 17. Telomere fusion can arise from Pol β overexpression.

In addition, two-fold overexpression of Pol β was found to promote genomic instability due to Pol β substitution for replicative and recombination pathways, processes that typically rely on high fidelity polymerases.[11] For instance, Pol β overexpression in Chinese hamster ovary (CHO) cells resulted in a 4-12 fold increase in mutational events in cells. These cells also exhibited reduced sensitivity to bifunctional DNA-alkylating agents (e.g. cisplatin) because Pol β bypassed the lesions during translesion synthesis.[11] Briefly, translesion synthesis (TLS) is a DNA damage tolerance process that allows DNA replication machinery to replicate past a lesion via TLS polymerase switching (Figure 18).[120] Together, the reduced sensitivity and increased genomic instability phenotype in cancer cells may promote cancer progression.

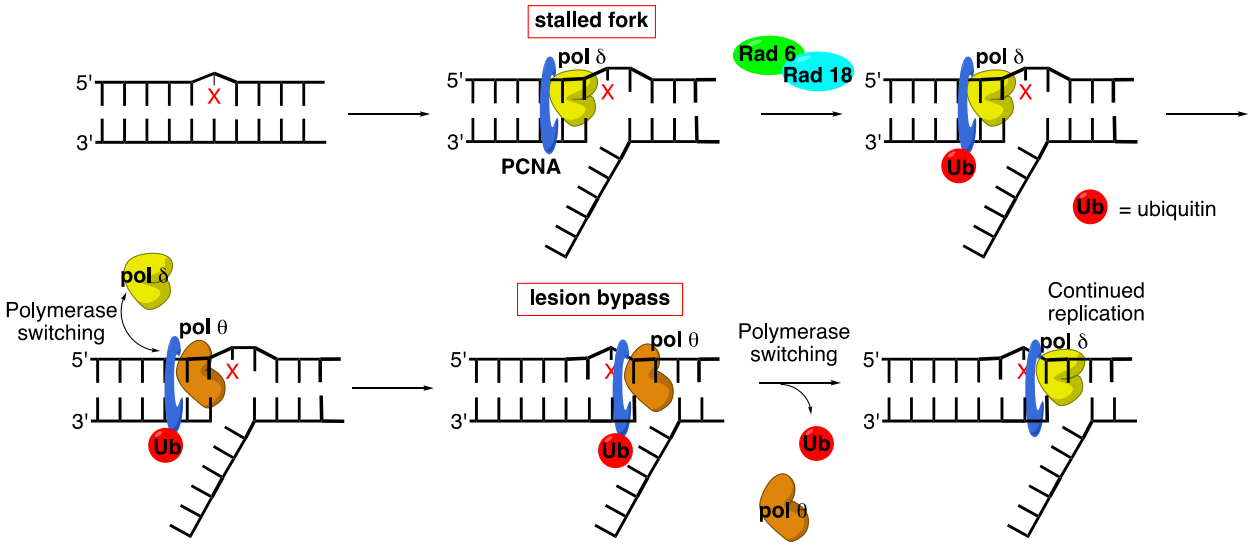


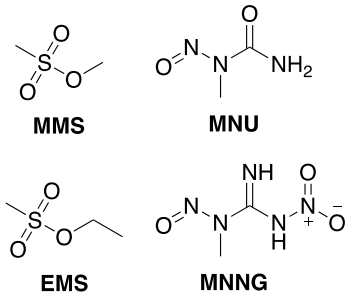
Figure 18. Simplified mechanism of translesion synthesis.

Pol β is constituitively expressed but is adaptive to cellular conditions. Expression of Pol β, and in turn BER, is up-regulated in response to oxidative stress.[121] For instance, oxidative stress induced by treatment with lipopolysaccharide (LPS) in mouse monocytes increases the expression of Pol β. This led to increased BER activity in cells. LPS treatment in Pol β null cells exhibited more DNA damage, consistent with the idea that oxidative stress decreased the accumulation of DNA lesions on account of elevated levels of Pol β-dependent BER.[121]

Many cancer cells exhibit higher levels of DNA damage and generation of ROS. Pol β is overexpressed in many cancer cell lines.[11] Overexpression is accompanied by decreased sensitivity to anticancer drugs, including cisplatin.[11] In fact, the Pol β overexpression was associated with tumor metastasis and poor prognosis in esophageal cancer patients.[122] This is likely to be disease-dependent considering that low Pol β levels were also associated with aggressive phenotype and poor prognosis in breast cancer.[123] Regardless, a decrease in BER due to Pol β inhibition precludes the reduced sensitivity to anticancer drugs. The reverse also holds true in that low levels of Pol β expression result in deficient BER and hypersensitivity to DNA-damaging agents.[4,14,19,78,124]

A decrease in Pol β activity was first analyzed by using its own subdomains that lack polymerase activity. A 14 kDa truncation of the enzyme consisting of the lyase domain and the DNA binding domain (Figure 15, Lyase and D subdomains) that is capable of binding ssDNA and dsDNA but incapable of catalyzing polymerization is a competitive inhibitor of the intact enzyme. This truncated enzyme competed with Pol β for the primer-template substrate, reduced Pol β activity, and decreased BER efficacy *in vitro*.[124] The 14 kDa domain inhibited Pol β activity to a greater extent than the Klenow Fragment and Pol α, two non-homologous polymerases, likely due to differences in substrate binding.[124] Pol β, as well as the 14 kDa truncation, acts on gapped DNA containing 3'-OH/5'-phosphate termini and this binding mode is unique to Pol β.[81]

Furthermore, Pol β knockout mice embryonic fibroblasts (MEFs) exhibited higher levels of cell death when exposed to monofunctional alkylating agents: methyl methanesulfonate (MMS), ethyl methanesulfonate (EMS), N-methyl-N-nitrosourea (MNU), and N-methyl-N-nitrosoguanidine (MNNG).[14] Furthermore, Pol β knockdown by RNA interference (RNAi) induced hypersensitivity to MMS in MEFs.[19] Pol β-deficient cells were not hypersensitive to bifunctional alkylating agents, such as cisplatin and nitrogen mustard, as these tend to form bulkier lesions that require NER instead of BER. In addition, a lack of Pol β also does not enhance IR cytotoxicity.[14,19]



Cell sensitivity to BLM in response to Pol β deficiency is inconsistent across several existing studies.[17,18,125] Knockdown of Pol β in MEFs exhibited BLM hypersensitivity and increased oxidative damage and gene mutations.[125] Pol β deficiency induced by triterpenoid inhibition resulted in the potentiation of BLM cytotoxicity.[18] Similarly, suppression of the lyase activity via plant sterol inhibitors produced the same result.[17] However, other studies report no significant BLM sensitivity in Pol β knockout MEFs and knockdown by RNAi.[14,19,126] Altogether,

there are more reports that Pol β inhibition enhances BLM sensitivity; however, the reason for conflicting results is still unclear.

Balanced levels of Pol β are important for genomic stability. Knockouts of Pol β, similar to other DNA repair enzymes, are embryonic lethal in mice.[127,128] Embryonic lethality arises when gene expression is required during embryogenesis. Lack of Pol β during embryonic development results in embryo death. However, knock down of Pol β in mature cells does not lead to cell death. Several active Pol β variants have been identified in patient tumors (Section 2.5.1).[13,129–131] The presence of cancer-specific mutants is significant because it presents a space for developing inhibitors that will preferrentially target mutants associated with a certain cancer over WT Pol β in healthy cells. If successful, this strategy could reduce non-specific toxicity associated with chemotherapeutics.

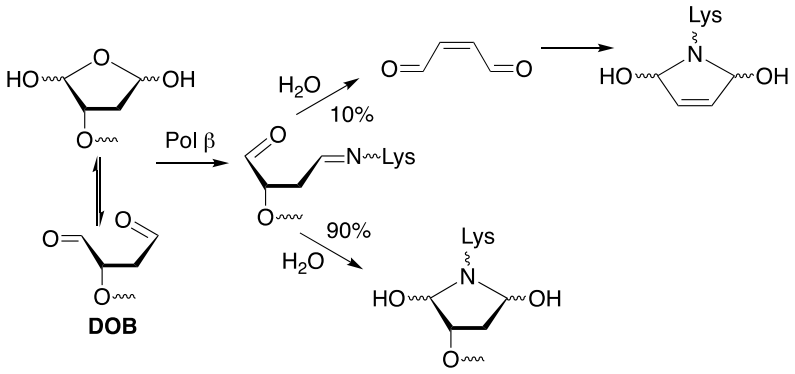
2.3.1.2 BER inhibition by oxidized AP lesions

Oxidized AP lesions are processed with varied success by BER enzymes.[30,91,132] Ape1 acts on DNA containing L and C4-AP lesions at a slower rate than incision of DNA containing AP.[90,91] In addition, Pol β is irreversibly inhibited by DOB and pC4-AP lesions by forming a covalent adduct with the active site Lys residue.[29,30]

As mentioned before, antitumor antibiotics, BLM and NCS, induce DNA damage via sugar oxidation and are very cytotoxic (Section 2.2.2.1). In an attempt to understand how these agents were so effective at killing cells, oligonucleotides containing oxidized AP lesions, DOB and pC4-AP, were synthesized and used as homogenous substrates for Pol β.[29] The more potent inactivator of Pol β, DOB covalently modified Lys72 and Lys84, consistent with inactivation by Schiff base formation between a active site Lys and the dialdehyde of DOB. With radiolabeled substrates, it was determined that 90% of the inhibition occurs from direct trapping by DOB and 10% occurs

from the release of butenedial that reacts with the enzyme (Scheme 9).[29] Product analysis of Pol β with DOB-containing DNA with and without FEN1 exhibits limited polymerase activity, suggesting DOB inhibits both SP-BER and LP-BER.[30] The inhibition of Pol β by oxidized AP lesions can lead to more severe biological consequences than decreased BER. For example, DOB inactivation of Pol β led to strand breaks in a trinucleotide repeat (TNR) tract, which has implications in neurodegenerative diseases.[133]

Scheme 9. Inactivation of Pol β by DOB.



Inhibition of Pol β by DNA containing these lesions are not selective. For instance, pC4-AP also inactivates Pol λ and Pol θ at the active site Lys residue in the respective lyase domains.[134,135] Oxidized AP lesion, DOB, also inactivates Pol λ and DNA Ligase I.[133,134] This is significant because Pol λ, whose primary role is in NHEJ to repair DSBs, acts as a back-up polymerase in BER. Inactivation of both Pol β and Pol λ may assist in toxicity to cells. In a similar way, the inhibition of both enzymes may increase the ability of DOB to decrease BER activity. In contrast, inhibition of multiple enzymes is generally undesired as it decreases selectivity and may induce adverse effects.

2.4 DNA Repair Inhibition

DNA repair is a crucial mechanism for drug resistance pertaining to drugs that damage DNA.[1] BER plays a role in resistance to alkylating agents by removing drug-induced DNA lesions. In addition, NER confers resistance against crosslinking agents, such as cisplatin. Importantly, tumor cells often have higher levels of certain DNA repair pathways to combat increased levels of DNA damage and maintain enough genomic integrity to avoid apoptosis. Reducing the effectiveness of repair pathways may capitalize on the difference between cancerous and healthy cells.[101] Therefore, repair inhibition is an attractive route to enhance cell sensitivity to DNA damaging agents.[14–19]

DNA repair inhibition is a growing field and therapeutic targets have been identified as possible routes to induce apoptosis in cancer cells as a monotherapy or to enhance the cytotoxicity of cancer therapeutics through combination therapy.[2,20,70] Noteworthy targets of BER inhibition include Ape1 and poly (ADP-ribose) polymerase (PARP).[2] Ape1 inhibition is unique to other targets because Ape1 also regulates reduction-oxidation activities in cells.[92] Small molecule inhibition of Ape1 enhances the cytotoxicity of DNA damaging agents, like MMS, H_2O_2, BLM, TMZ, cisplatin, and IR.[2] While some Ape1 inhibitors are nonspecific BER inhibitors, others are specific to Ape1 and are in various phases of clinical trials.[2]

One of the most successful examples of DNA repair inhibitors are PARP inhibitors (PARPi), such as olaparib or rucaparib, in the treatment of breast and ovarian cancer (Figure 19).[136,137] PARP is involved in repair through the recognition of SSBs, synthesis of poly-ADP ribose (PAR), and recruitment of other enzymes to initiate DNA repair.[138] PARPi trap PARP on the DNA and form PARP-DNA complexes. PARP-DNA complexes block the progression of the replication fork during DNA synthesis and lead to the formation of DSBs.[139]

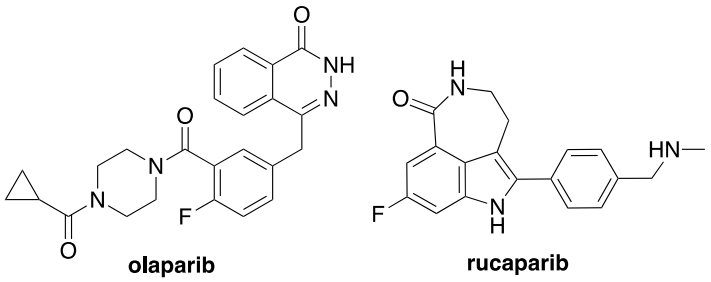


Figure 19. Structures of PARPi.

PARPi take advantage of an interaction called synthetic lethality. Two genes are considered synthetic lethal when a deficiency in both leads to cell death, whereas a deficiency in one of those genes does not. Treatment with a PARPi is most effective in patients with the BRCA1-mutant, a gene that encodes a tumor suppressor protein that functions in HR.[139,140] The BRCA1 mutation in cells results in a deficiency in HR DSB repair and the cell resorts to error-prone NHEJ.[2] More clearly, PARP inhibition is not toxic to healthy cells that are proficient in HR. However, the loss of both HR and PARP activity in cancer cells results in cell death. The PARP-DNA complex formed as a result of PARPi is more cytotoxic than unrepaired SSBs because the repair of the PARP-DNA complexes requires multiple enzymatic pathways, including Pol β.[139] Pol β knockouts are associated with PARPi hypersensitivity, suggesting that Pol β inhibitors and PARPi may work together to promote cell death.[15] Furthermore, Pol β has been postulated to be synthetic lethal in homologous recombination deficient cells (e.g. BRCA1/2).[12]

2.5 Pol β is an anti-cancer target

Pol β has been identified as an attractive therapeutic target as a means to enhance the cytotoxicity of DNA damaging agents.[2,20,21,70,141] Pol β is a good candidate for multiple reasons: it is overexpressed in cancer cells, its inhibition has improved the efficacy of chemotherapeutics, several variants have been identified in tumors, and it is the primary polymerase in BER. Through

the specific and efficient inhibition of Pol β in cancer cells exposed to DNA damaging agents, the accumulation of DNA damage induces apoptosis in cells (Scheme 1).

As mentioned before, the balanced regulation of Pol β is necessary for proper genomic maintenance (Section 2.3.1.1.3). However, Pol β is overexpressed in many cancer cell lines, including ovarian, colon, and leukemia.[11] Overexpression is accompanied with decreased sensitivity to anticancer drugs.[11] The cancer-specific (in contrast to healthy cells) overexpression suggests Pol β activity is more important for cancer cell viability and its inhibition may be more toxic to cancer cells than in healthy cells. In contrast, low levels of Pol β expression result in deficient BER and hypersensitivity to DNA damaging agents.[14,19,124] In support of this, inhibition of Pol β enhance the efficacy of chemotherapeutics, such as monofunctional alkylating agents.

2.5.1 Pol β variants

There is an abundance and variety of Pol β variants in patient tumors.[13,129–131] In fact, Pol β mutants were identified in 30% of tumors in different tissues (e.g. colon).[13] Variants associated with other cancers have also been identified (e.g. gastric and ovarian).[130,131] For instance, a T889C point mutation in the coding region of the gene encoding Pol β, POLB, in gastric tumors resulted in mutant L259S and correlates with gene overexpression.[130] In colorectal cancer, 40% of patient tumors harbor mutations in the POLB gene.[13] Of these, 31% of mutations result in an amino acid substitution that can alter protein structure and function. Mutations were found in both domains of the enzyme and the effect on enzymatic function varied depending on the identity and location of the mutant.[13,129]

The impact on function was difficult to predict via computational algorithms. For example, lyase domain mutant T79I was predicted to have decreased function but *in vitro* enzyme activity was normal relative to WT.[13] Alternatively, G231D mutant in the polymerase domain was

predicted to be tolerated but the enzyme activity was significantly reduced and resulted in chromosomal aberration phenotype, specifically a missing, extra, or irregular portion of chromosomal DNA.[13,129] Functional Pol β mutants that are expressed in tumor cells represent a space to develop selectivity for cancerous cells over healthy cells. This has the potential to reduce non-specific toxicity.

2.6 Pol β inhibitors

Reported small molecule Pol β inhibitors include bile acids, sulfolipids, naphthoic acids, and others.[20,21] Currently, there are no Pol β inhibitors in the clinic because inhibitors identified often lack defined selectivity and binding region. In addition, several identified inhibitors exhibit poor potency (low μM). The potencies of these compounds are typically reported as half maximal inhibitor concentration (IC_{50}) values, which is a measurement of the effectiveness of a substance and equates the amount required to inhibit half of a specific enzyme (in this case, Pol β). In some cases, the efficiency of an inhibitor is reported as dissociation constant (K_D), which is the concentration of a ligand needed to occupy 50% of receptors.

2.6.1 Small molecule inhibitors

The first small molecule inhibitor(s) of Pol β identified were fatty acids, linoleic acid (LA) and nervonic acid (NA) (Figure 20).[22,142,143] LA and NA were shown to bind the 8 kDa lyase domain and have IC_{50} values of 38 and 5.8 μM, respectively. However, these inhibitors are not selective because they also reduce the activity of Pol α, a non-homologous polymerase involved in the initiation of DNA replication.

Other Pol β inhibitors are natural products: koetjapic acid (KA) and pamoic acid (PA) (Figure 20).[20,144,145] These also bind to the 8 kDa domain of Pol β with varying inhibitory efficiencies. KA has a reported IC_{50} value of 20 μM and PA has a K_D of 9 μM.[144–146] While PA

binds the lyase domain, it inhibits both the lyase and polymerase activity of Pol β and was shown to enhance the efficacy of alkylating agent, MMS.[144,145]

In addition, lithocholic acid is a bile acid produced in liver cells that is active against Pol β (Figure 20).[147] This compound binds the 8 kDa domain and exhibits an IC_{50} = 11 μM and K_D = 1.56 mM. Lithocholic acid suppresses both lyase and polymerase activity and inhibits Pol β and α.[147] In addition to the nonselective inhibition of Pol β, lithocholic acid interacts with several other cellular processes, including p53 degradation, inflammatory reactions, generation of ROS and reactive nitrogen species, which promotes colon carcinogenesis.[148]

More recently, honokiol (HKL), a biphenol compound, was identified as a Pol β inhibitor (Figure 20). HKL inhibits both Pol β (IC_{50} = 4.0 μM) and Pol λ (IC_{50} = 8.3 μM).[25] The inhibition of both Pol β and Pol λ is significant because this inhibitor can inactivate the primary and back-up polymerase of BER. HKL inhibits polymerase activity and enhances BLM cytotoxicity in three different cancer cell lines.[25]

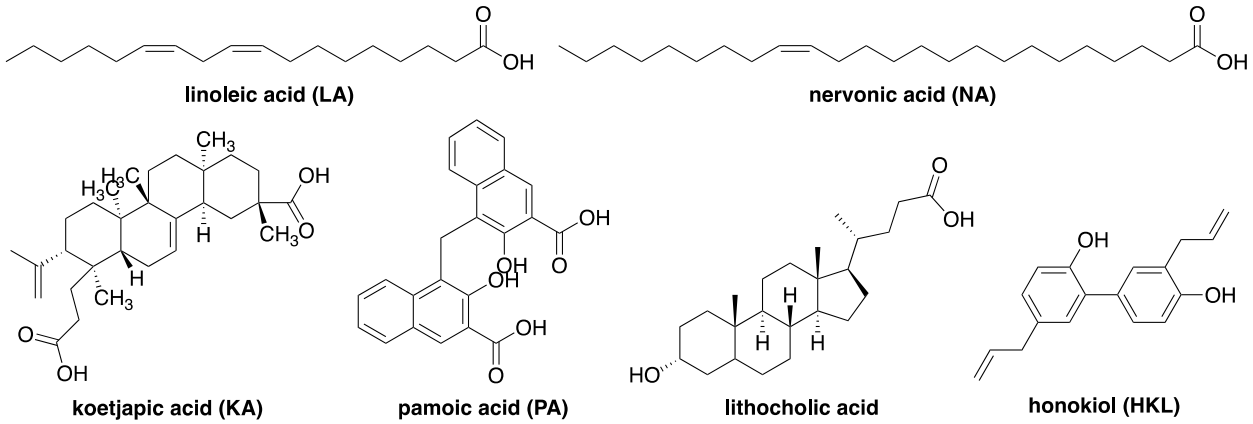


Figure 20. Structures of natural product Pol β inhibitors.

As mentioned before, oxidized AP lesions in DNA inactivate Pol β (Section 2.3.1.2).[29] This discovery inspired the design of a small molecule inhibitor library (**1**) of dinucleotide analogs that mimic the DOB lesion (Figure 21).[26] The dinucleotide analog contained a methylene between the DOB moiety and the phosphate to improve chemical stability and discourage elimination.

Structural diversity was introduced at the 3' end of the molecule through the introduction of an oxime that was diversified using a library of aldehydes. Candidate **2** (IC_{50} = 21 μM) irreversibly inhibits Pol β *in vitro* (Figure 21). A bis-acetate pro-inhibitor (**pro-2**) was more effective in cell lysates and more cytotoxic.[26] **Pro-2** (20 μM) potentiates the cytotoxicity of MMS between 2- and 5- fold in prostate cancer cells, depending on the alkylating agent concentration. The pro-inhibitor is specific for Pol β over Pol λ and the Klenow fragment.[26] Inhibitor **pro-2** is the first known irreversible inhibitor of Pol β and its synergistic effect with MMS supports the proposal of using a DNA damaging agent simultaneously with a DNA repair inhibitor as a viable cancer therapy.

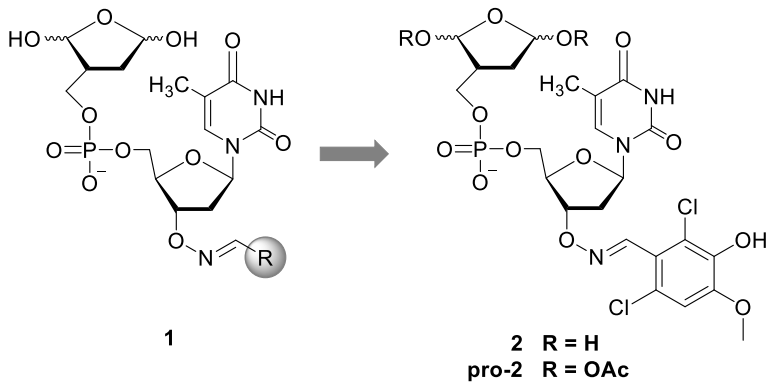
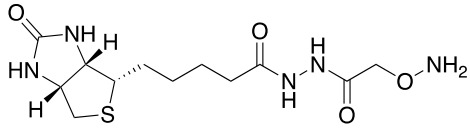


Figure 21. Irreversible inhibitor library (**1**) led to identification of inhibitor (**2**).

A second generation inhibitor stemming from the previous study was designed to further functionalize **pro-2** at the 5-methyl position of the thymidine nucleobase to probe other regions of the enzyme.[27] Analysis of this new library (**3**) led to identification of **4** as a 50-fold more potent Pol β inhibitor (IC_{50} = 0.4 μM) than previously studied **2** (Figure 22). Compound **4** also inhibits Pol λ (IC_{50} = 0.25 μM), which may enhance toxicity by inactivating both BER polymerases. The overall potentiation of MMS cytotoxicity was greater than that of BLM when **pro-4** was used in cervical cancer HeLa cells.[27] Levels of unrepaired dRP groups in the DNA were examined using an aldehyde reactive probe (ARP) in response to MMS (0.3-0.4 mM) and **pro-4** (5 μM). A 3-fold increase of ARP


aldehyde reactive probe (ARP)

reactive sites in cells was observed after the removal of the damaging agent, suggesting that **pro-4** prolongs the lifetime of BER intermediates in cells (i.e. SSB with a 5'-dRP).[27] When cells were treated with MMS and **pro-4** for 2 h, followed by media replacement containing only **pro-4**, cell death continued to increase.[27] DNA repair inhibition persists even after removal of the DNA damaging agent. This is consistent with the ability of **pro-4** to prevent repair and extend the lifetime of DNA lesions that induce cell death.

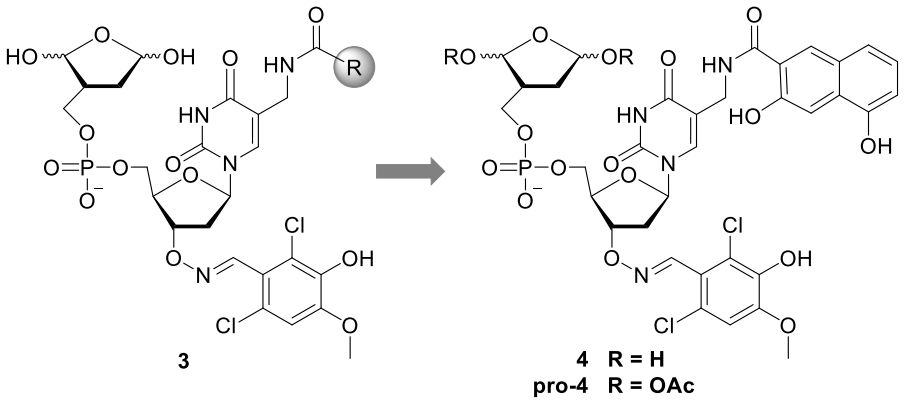


Figure 22. Second generation of irreversible Pol β inhibitor.

Statement of the project

The goal of this project was to identify potent inhibitors selective for Pol β that are useful in cells and improve the selectivity of cancer treatment either by the efficacy of DNA damaging agents, such as MMS and/or BLM, or a synthetic lethal interaction between Pol β and repair-deficient cancer cells.

3. Results & Discussion

3.1 First-generation library scaffold design

Our goal was to improve the potency and selectivity of Pol β inhibitors by combining an electrophile and a molecule that resembles DNA with structural diversity. These molecules were designed based on how DOB inactivates Pol β (Section 2.3.1.2). While inhibitors have been synthesized previously using this approach (**2**, Figure 21 and **4**, Figure 22), we were not convinced that earlier strategies were the only way to generate inhibitors.[26,27] We, purposefully, tried to avoid previous approaches due to the low yielded synthesis of the C5-aminomethyl group, necessary for functionalization at that position.[27] Prior to jumping into various synthetic routes, several issues were considered.

A necessary aspect of an inhibitor library is generating structural diversity, which is crucial for identifying a potent and selective molecule. Introducing a variety of groups at a specific location on the scaffold allows for unique interactions between Pol β and the library molecule. First-generation inhibitors can be enhanced by introducing additional structural diversity at a different location on the molecule, resulting in a second-generation compound (Figure 23). However, we were unsure of the best method to introduce diversity, as well as the order of introducing structurally unique groups (X and Y in Figure 23).

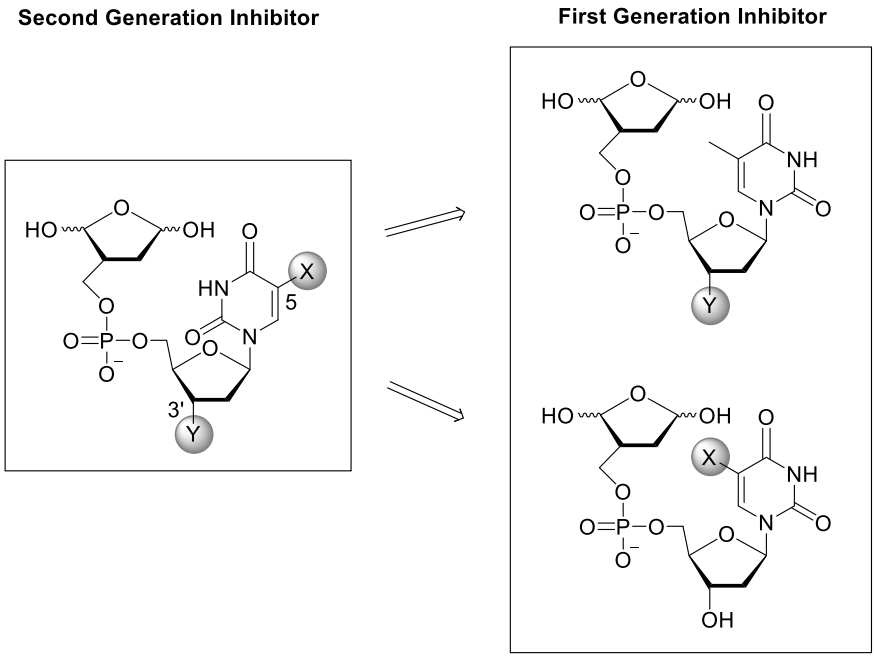
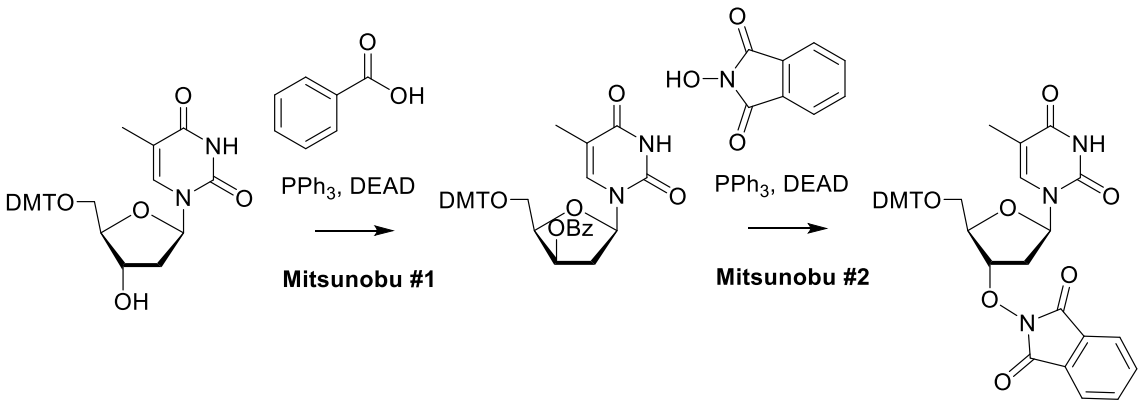


Figure 23. Retrosynthetic design of second-generation inhibitor library.

For our scaffold, the two most rational locations to introduce structural diversity are the 5-methyl position of the thymidine base (X, Figure 23) and the 3'-position of the sugar (Y, Figure 23). These sites require the least modification of the general structure (i.e. dinucleotide containing DOB). Reactions at these positions have already been established, negating any need for the development of extensive "new" chemistry. In a previous library, structural diversity was introduced by a 3'-alkoxyamine that was coupled to a variety of aldehydes.[26] The synthesis of the precursor for preparing the oxime library required consecutive Mitsunobu reactions to orient the alkyoxyamine with the correct stereochemistry at the 3'-position (Scheme 10). Therefore, we wanted to explore other options for diversification to improve the library synthesis. The inhibitor identified in the oxime library (**2**, Figure 21) was further functionalized at the 5-methyl position of thymidine through an amide bond (**4**, Figure 22).[27] In the second-generation library (**3**, Figure 22), an azide was incorporated at the C5-methyl group of thymidine and following reduction, the amine scaffold was reacted with a library of carboxylic acids (**3**).

Scheme 10. Consecutive Mitsunobu reactions for alkoxyamine scaffold.



Structural diversity is introduced at these positions by incorporating a suitably reactive functional group in the scaffold at either of these positions. Functionalization via reductive amination or amide coupling are attractive routes for making libraries. These transformations are well established for preparing libraries.[26,27,149,150] Depending on the polarity of the reaction with the scaffold, a library of small molecules containing either an aldehyde, a carboxylic acid, or an amine would be useful. For example, a scaffold bearing an amine would react with a library of carboxylic acids to form amides. Fortunately, small molecules bearing these groups are vast, commercially available, and structurally diverse. These qualities are attractive for generating a library.

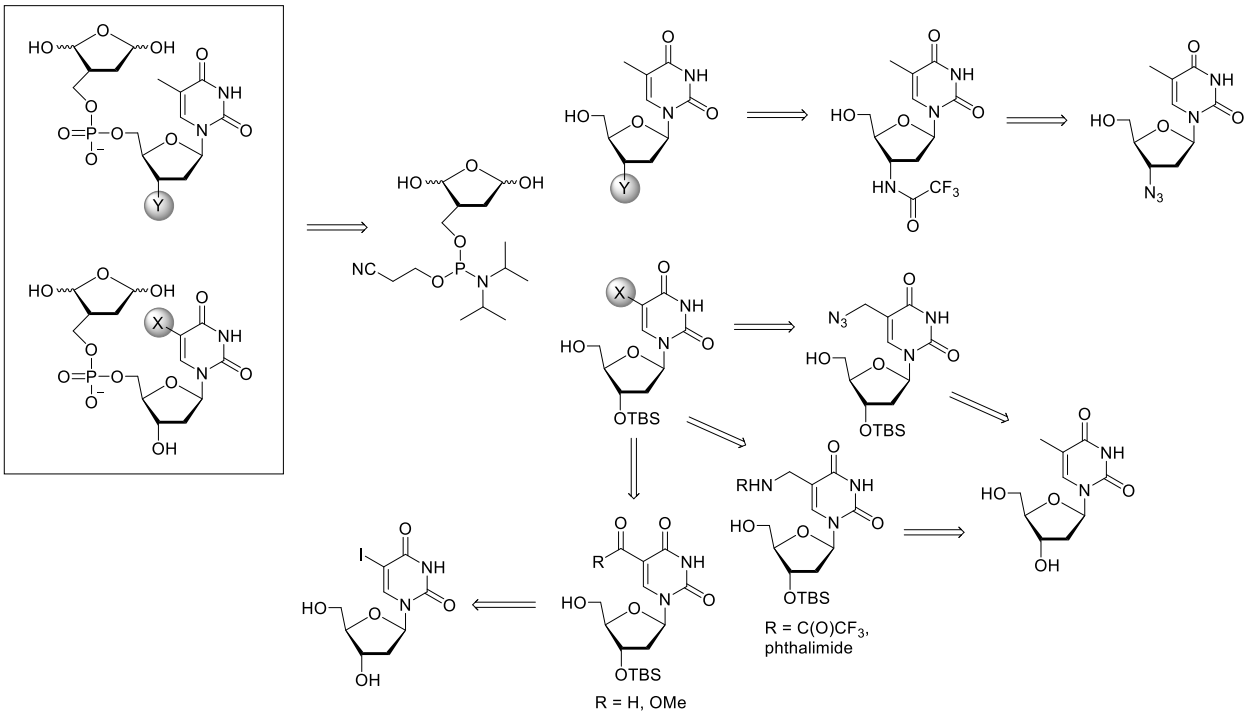
An additional question we had was regarding the best way to disguise the electrophile that reacts with the enzyme. The DOB portion is an electrophilic moiety and, therefore, reactive with nucleophiles. The 1,4-dialdehyde has been used to trap lysine residues before.[151] The electrophile (acyclic 1,4-dialdehyde) exists in equilibrium with its cyclic DOB form, a disguised version of the electrophile.[30] The aldehydes must be protected for the entirety of the synthesis to prevent undesired reactions until the very last step when they are removed to reveal the DOB. Originally, pentenoyl groups were chosen because we needed a mild, quick method for cleaving acetals.

Pentenoyl groups have a long-standing history as aldehyde protecting groups, particularly in carbohydrate chemistry.[152–155] Pentenoyl protecting groups are stable in basic conditions and are compatible with common protecting groups, such as silyl and acetamide groups required in the synthesis of the scaffold molecule. They are most often cleaved under mild oxidative conditions with *N*-bromosuccinimide (NBS). However, pentenoyl groups are not the only groups that can be used to mask the DOB. For instance, aldehydes can also be protected with acetate (Ac) or nitroveratryl (Nv) groups.

Additional synthetic considerations involved the nucleobase and phosphate linkage used to mimic the DNA structure. Pol β does not exhibit a preference for any of the natural nucleobases. Therefore, thymidine was included in the scaffold because it is the easiest to work with (i.e. no protecting groups required on the nucleobase during synthesis) (Figure 23). An additional methylene between the phosphate and the 5'-DOB was included in the design to improve stability by discouraging elimination.

As mentioned before, structural diversity is a necessary aspect of generating a library. At the 3'-terminus of the molecule, we chose to mask an amine via a trifluoroacetamide protecting group. However, we were uncertain about the functional group at the 5-methyl position, as well as the method for diversification (e.g. reductive amination or amide coupling). Therefore, several approaches for synthesizing a first-generation library were explored (Scheme 11).

Scheme 11. Retrosynthetic pathway of first-generation inhibitors(s).



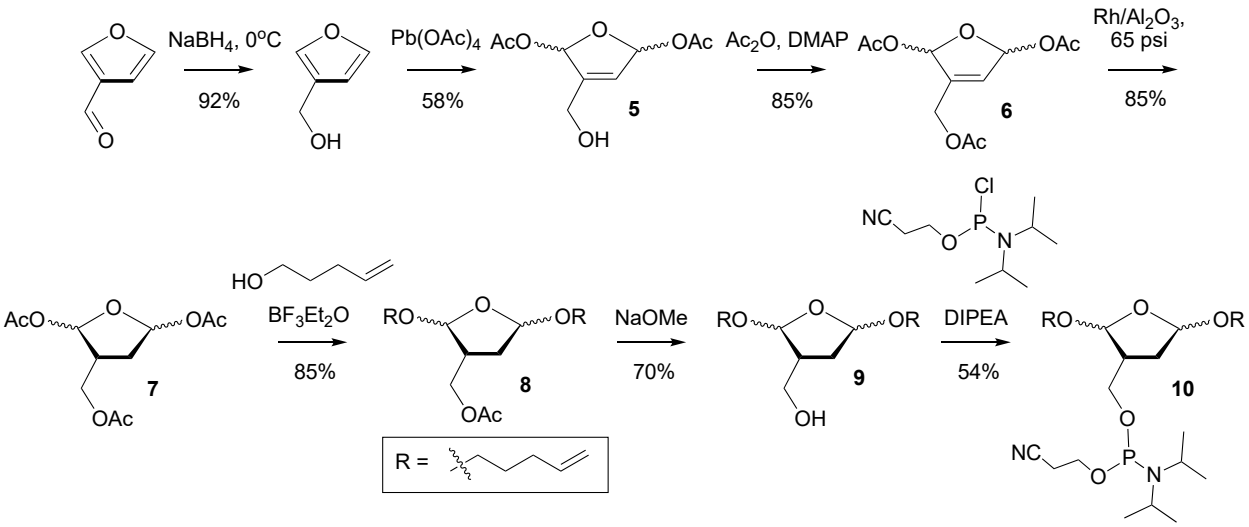
In an attempt to prepare a library via reductive amination, we tried to synthesize a scaffold containing a C5-aldehyde. However, generating a library via reductive amination between this scaffold and several amines proved to be difficult due to the instability of the products (Section 3.1.1). Alternatively, we tried to synthesize a library via amide coupling. First, we tried to make a scaffold containing a C5-carboxylic acid and reacting it with a library of amines (Section 3.1.2). Ultimately, this was unsuccessful, so we chose to reverse the polarity of the amide bond formation by coupling a library of carboxylic acids with a scaffold containing a C5-amine (Section 3.1.2). After multiple iterations of synthetic routes (Schemes 19-22), we determined that preparing a C5-aminomethyl group protected as its *N*-trifluoroacetamide proved to be most useful for generating a first-generation library (Section 3.3.1).

This strategy of using a scaffold amine and carboxylic acid library was the same as previous second-generation inhibitor **4** (Figure 22), despite our hesitations regarding the approach (i.e. low

yield of C5-aminomethyl group).[27] Ultimately, we functionalized both diversification points (X and Y, Figure 23) through amide coupling. In retrospect, this was economically favorable because it allowed us to use the same set of carboxylic acids in all synthesized libraries.

Regardless of the nucleoside modification, the synthesis of the protected DOB moiety was constant (Scheme 12). The $NaBH_4$ reduction of 3-furanaldehyde, followed by acetal introduction by lead tetraacetate yielded intermediate **5**. This intermediate was acetylated (**6**) and hydrogenated (**7**). Reaction of **7** with pent-4-en-1-ol in the presence of a Lewis acid led to the introduction of the pentenoyl protecting groups (**8**). Subsequent deacetylation (**9**) and phosphitylation provided phosphoramidite **10**.

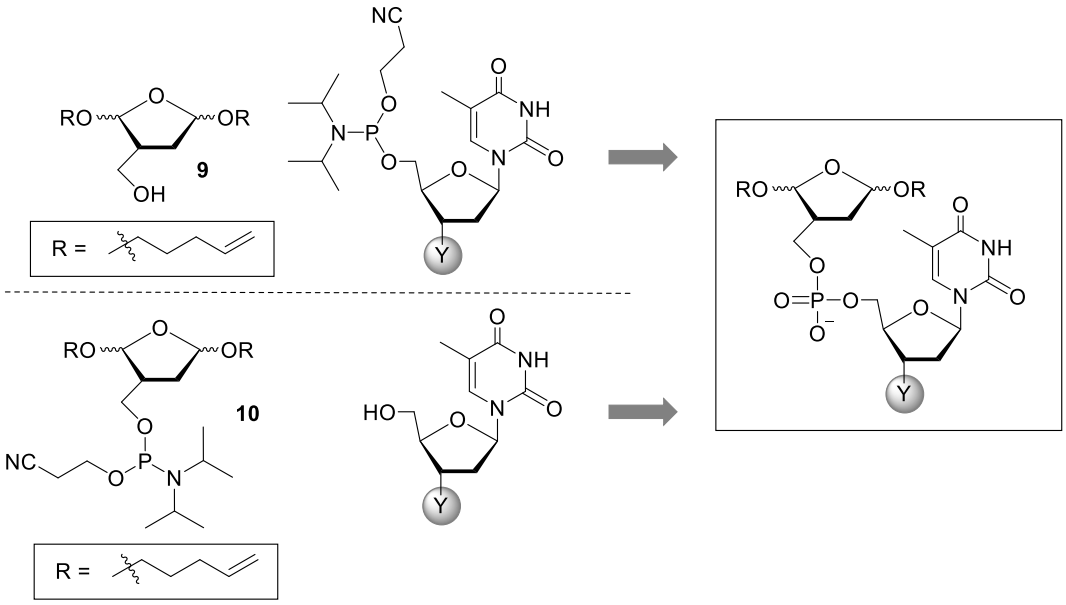
Scheme 12. Synthesis of DOB-portion of first-generation inhibitor.



The phosphoramidite was coupled to an alcohol to form the dinucleotide scaffold of the first-generation library. The polarity of this reaction can be carried out in either direction (Scheme 13). In some syntheses the phosphoramidite was prepared on the nucleoside (Scheme 13a) and in others it was part of the DOB portion (e.g. **10**, Scheme 13b). The direction of the polarity used depended on yield and usually required trying both directions to compare which resulted in a

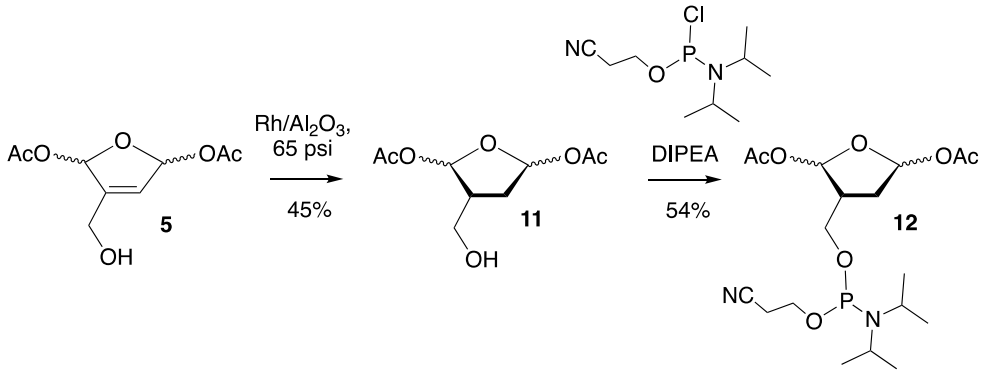
higher yield. It was difficult to predict which route would offer better results as it depended upon the thymidine derivative used.

Scheme 13. Phosphoramidite coupling in either polarity. (a) 3'-OH (**9**) and 5'-phosphoramidite or (b) 3'-phosphoramidite (**10**) and 5'-OH.



In some instances, the pentenoyl group was replaced by an acetate protecting group (Scheme 14). In this case, intermediate **5** was reduced (**11**) and phosphitylated to the acetate-containing phosphoramidite (**12**).

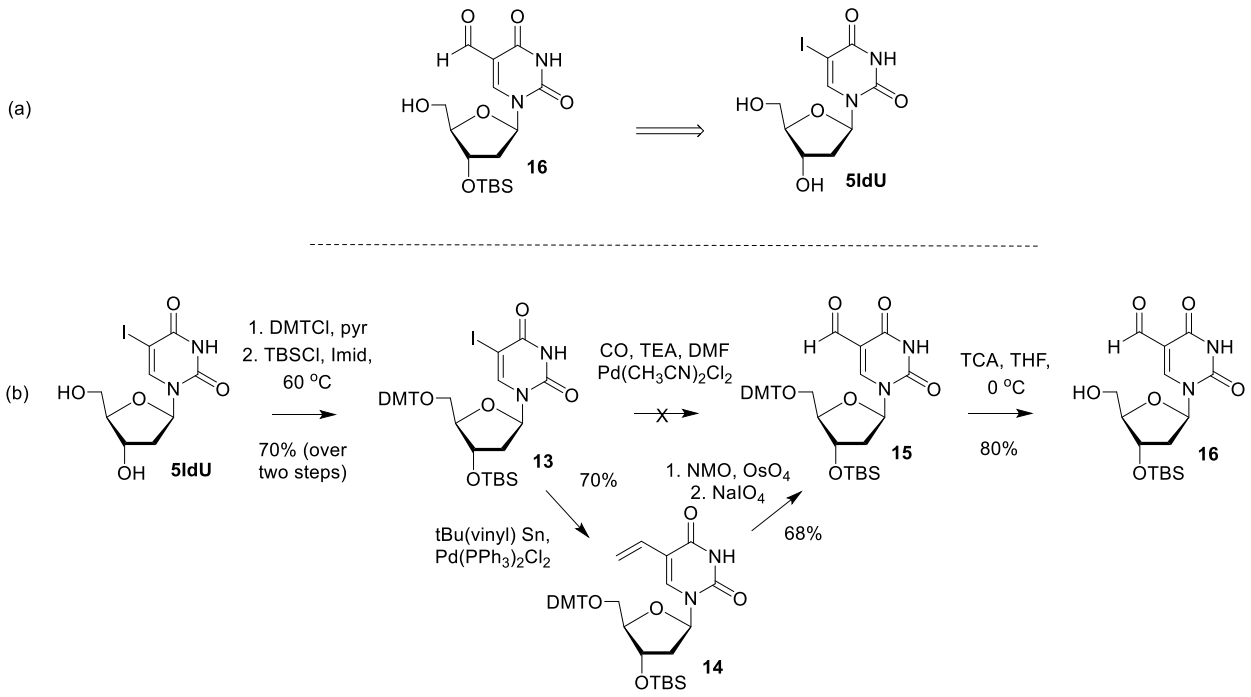
Scheme 14. Synthesis of the acetate-protected DOB phosphoramidite (**12**).



3.1.1 Diversification by Reductive Amination

Introduction of a carbonyl group on the nucleobase via an aldehyde started with 5-iododeoxyuridine (5IdU, Scheme 15). Following orthogonal protection of the 5'- and 3'-alcohols (**13**), the aldehyde was incorporated at the 5-position of **13** to yield **15**. Direct carbonylation was unsuccessful but vinylation (**14**) followed by oxidative cleavage led to the isolation of **15**. The 5'-dimethoxytrityl (DMT) group was selectively removed using trichloroacetic acid (0.1 M) to reveal **16**.

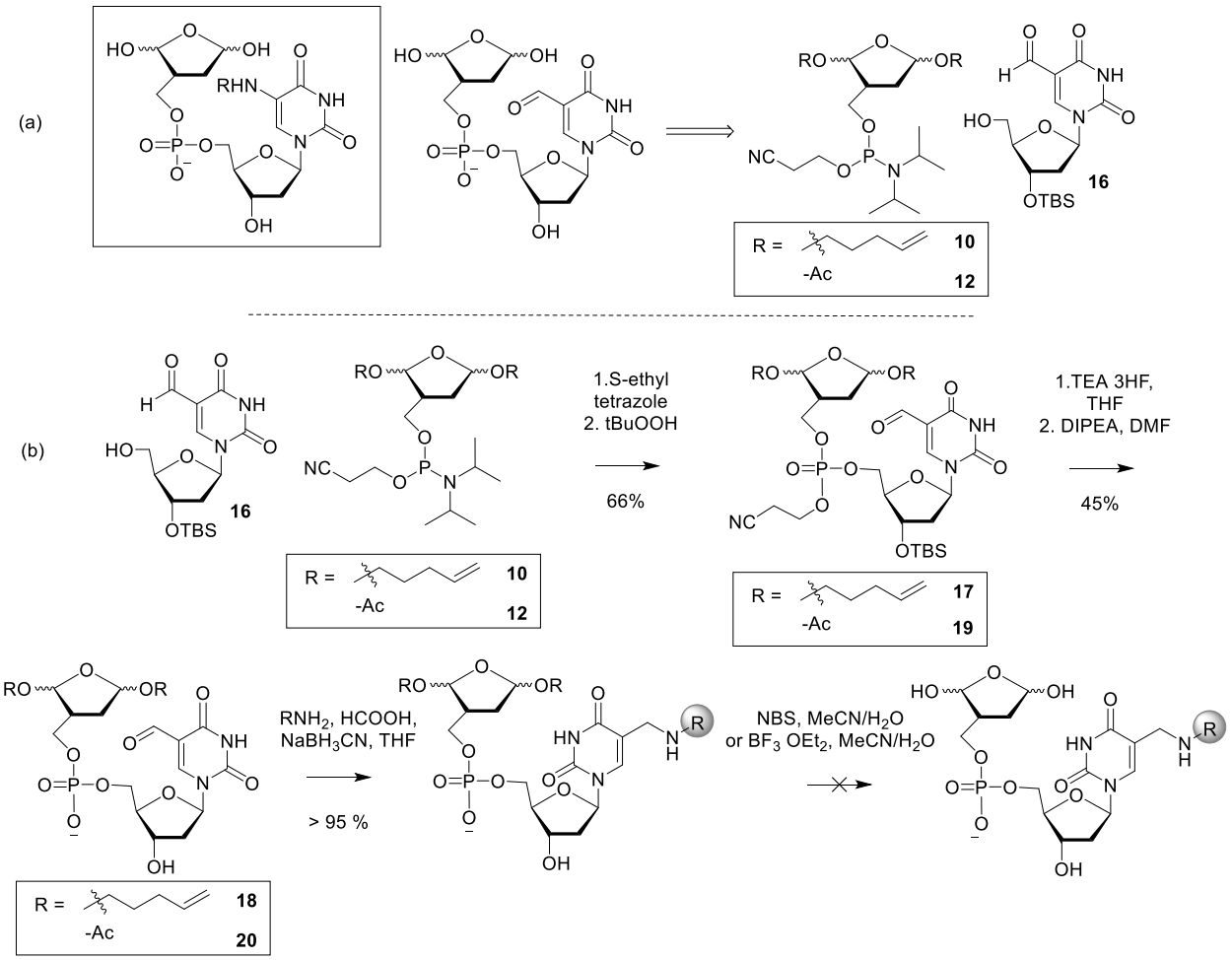
Scheme 15. Retrosynthetic (a) and forward (b) synthesis of aldehyde-containing nucleoside (**16**).



Compounds **10** and **16** were coupled to form the phosphate triester (**17**), which was deprotected to provide the scaffold (**18**, Scheme 16). Different amines (i.e. *n*-butyl amine, benzyl amine, and aniline) were coupled to the aldehyde via reductive amination. The reductive amination was completed on a small scale (< 500 nmol) and was confirmed by MS. However, NBS deprotection of the pentenoyl protecting groups on the DOB moiety led to oxidation of the secondary amine.

This scaffold was redesigned and resynthesized to bear acetate protecting groups instead of pentenoyl groups at the DOB moiety (Scheme 14). Phosphoramidite coupling of **12** and **16** provided phosphate triester **19** (Scheme 16b). Deprotection of the phosphate and the 3'-silyl group provided the acetate protected scaffold (**20**). Similar to the previous synthesis, reductive amination with model compounds (i.e. *n*-butyl amine, benzyl amine, and aniline) led to formation of a secondary amine. However, the acetate deprotection of the DOB did not yield the desired product. We speculate this was due to reactivity between the amine and the electrophilic DOB moiety. Based on this, we determined reductive amination was not a viable method for generating a first-generation library.

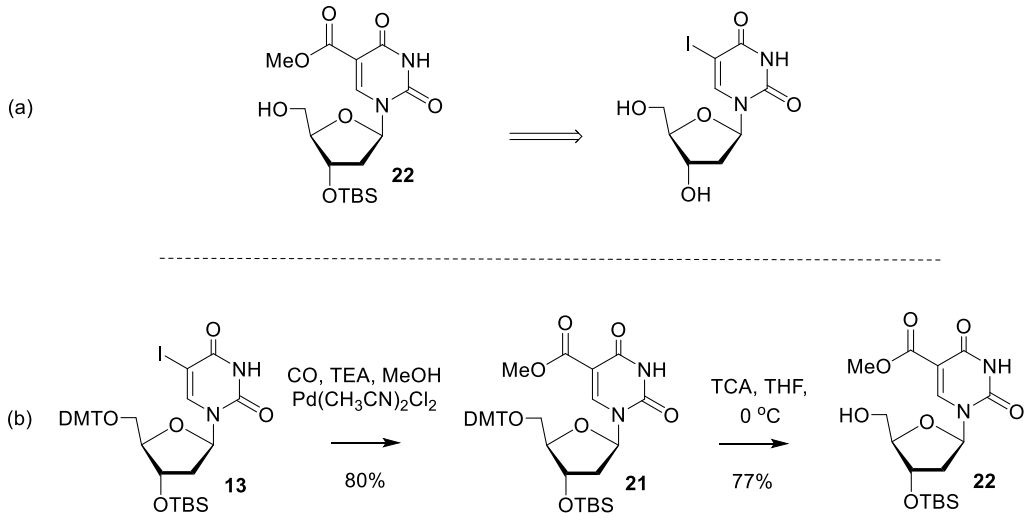
Scheme 16. Retrosynthetic (a) and forward (b) synthesis of aldehyde-containing scaffold **20**.



3.1.2 Amide Coupling Functionalization

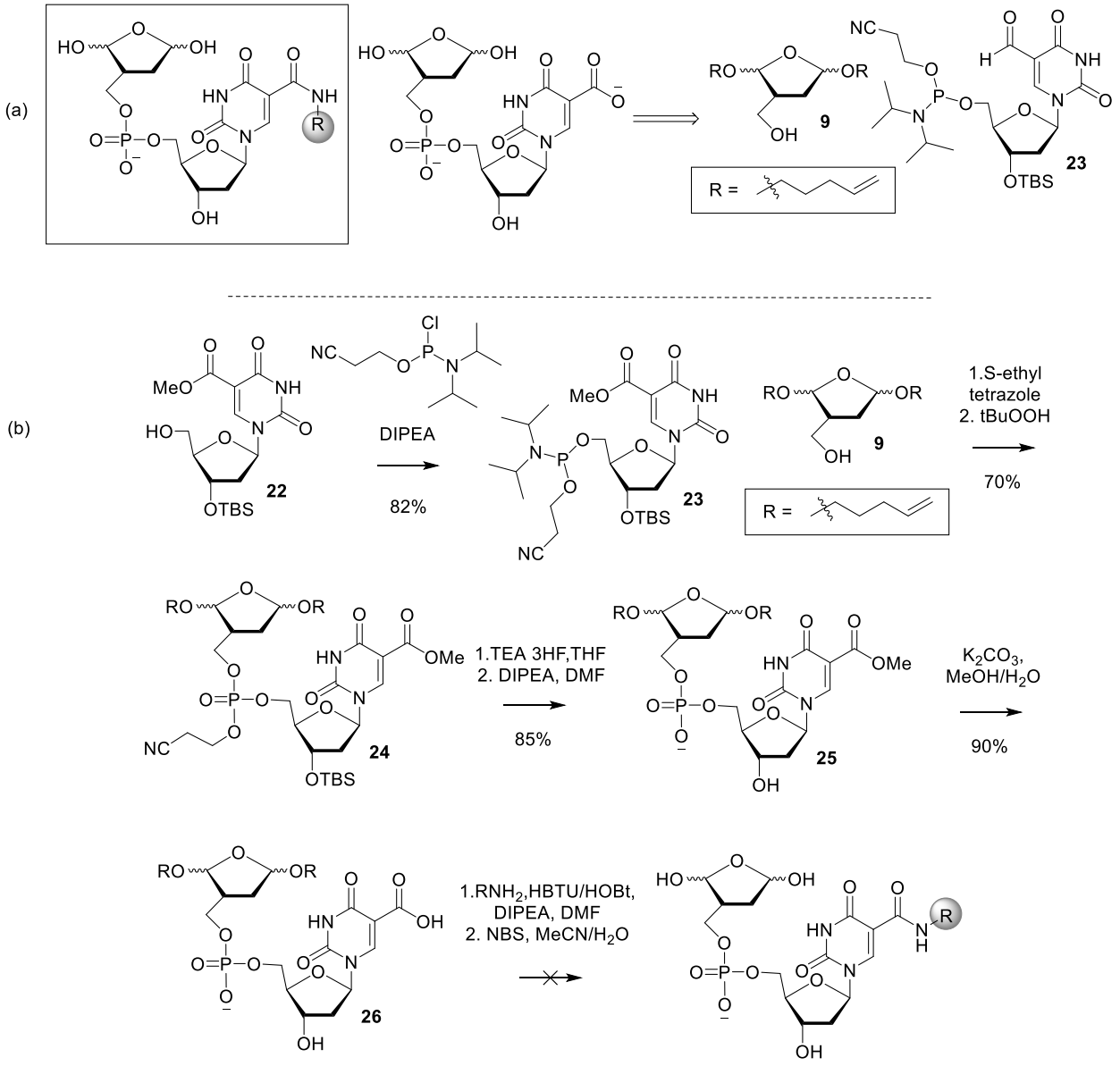
To avoid the generation of a secondary amine, a carboxylate at the 5-methyl position was introduced (Scheme 17). We hypothesized that the carboxylate could be functionalized by a library of amines to make amides that would be stable to the oxidation conditions used to reveal DOB and not react with this functional group after it is deprotected. Furthermore, in terms of synthesizing a nucleoside bearing a carboxylic acid, a portion of the synthesis was already completed. Previously synthesized **13** (Scheme 15) was carbonylated in the presence of methanol to form the methyl ester (**21**, Scheme 17b). Deprotection of the 5'-hydroxyl group yielded **22**.

Scheme 17. Retrosynthetic (a) and forward (b) synthesis of methyl ester-containing nucleobase **22**.



The 5'-hydroxyl group of **22** was phosphitylated (**23**) and coupled to previously synthesized **9** (Scheme 12) to yield phosphate triester **24** (Scheme 18). Following phosphate and 3'-hydroxyl deprotection (**25**), the methyl ester was cleaved under basic conditions to yield scaffold molecule **26**. Amide coupling of this scaffold was unsuccessful because the activating agents (i.e. HBTU and HOBt) reacted with the phosphate diester group. The activation of the phosphate resulted in coupling with the amine at both the phosphate and the carboxylate. Double coupling of the amine was observed by mass spectrometry and ^{1}H NMR. In addition, the chemical shift of the phosphate resembled that of a phosphoramidate triester ($\delta \sim 2$ ppm) in the ^{31}P NMR, which was further confirmation that the amine coupled to the phosphate. We hypothesized that the reverse polarity of the amide coupling (i.e. amine scaffold and carboxylic acid library) would be more successful and pre-activation of the carboxylic acid would prevent undesired activation of the phosphate group. This polarity is the same used in the synthesis of the previous second-generation inhibitor (**4**, Figure 22, Section 3.1).[27]

Scheme 18. Retrosynthetic (a) and forward (b) synthesis of carboxylic acid-containing scaffold **26**.

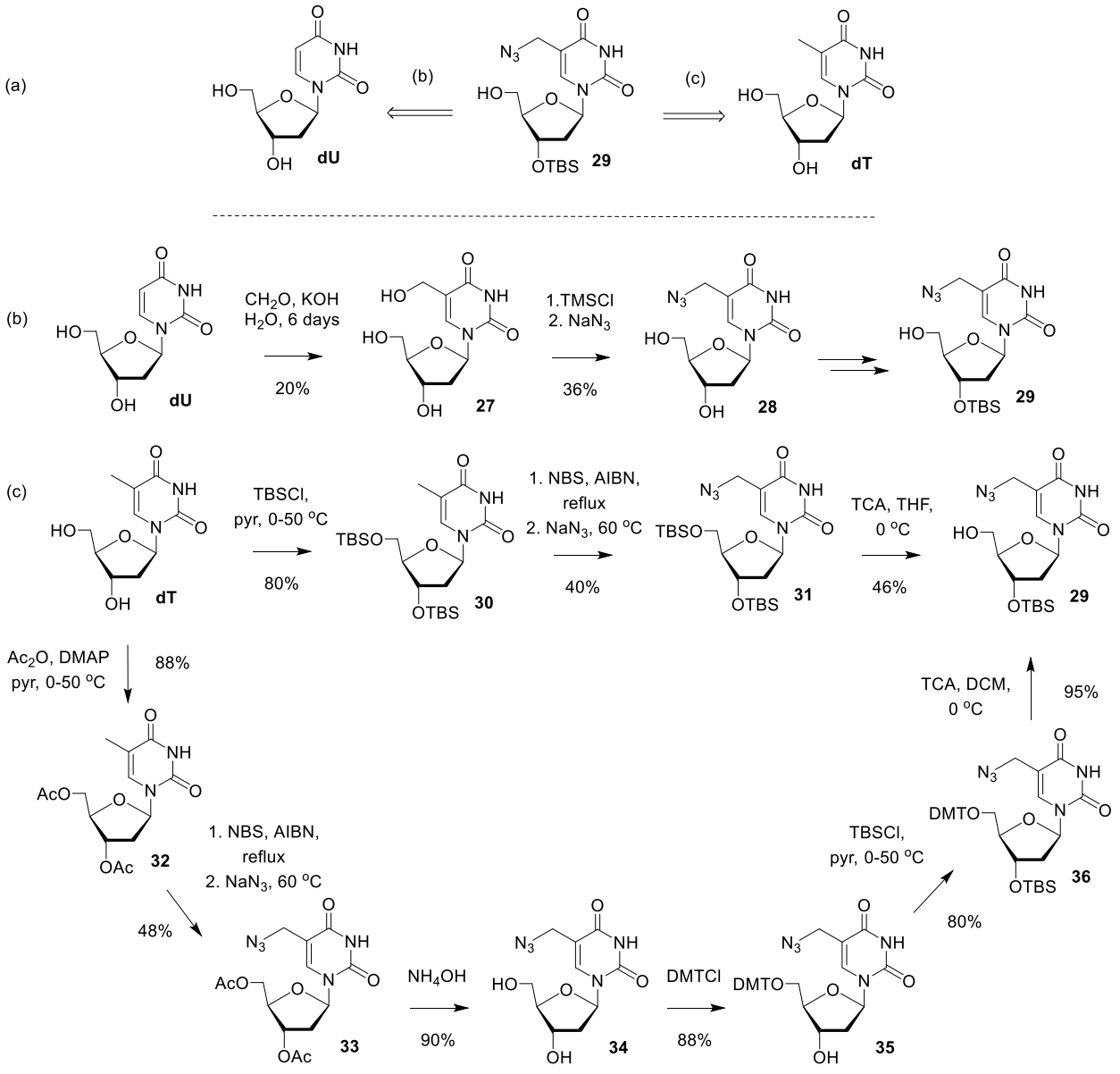


While planning the most efficient synthesis of the new scaffold (Scheme 19), we were aware of the challenges we were likely to face by starting from thymidine because of the difficulties in the synthesis of the previous inhibitor (**4**, Scheme 22).[27] Therefore, we tried a more direct approach that introduced the azide via the oxidation of dU (**27**), followed by alcohol

activation and substitution (**28**, Scheme 19b).[156] However, these first two steps required several days and were low yielding, so this pathway was not pursued further.

Therefore, we reverted to thymidine as a starting point and introduced the azide at the 5-methyl position (Scheme 19c). There are multiple ways in which this can be done. One method for introducing an azide at the 5-methyl position of thymine is via the bromide produced by the Wohl-Ziegler reaction.[157] Thymidine silylation (**30**) was high yielding and scalable to several grams. However, the following three steps involving one pot bromination and nucleophilic substitution by sodium azide (**31**) and subsequent selective 5'-desilylation to **29** were low yielding, less than 20% overall (Scheme 19c, top). When dT was protected with acetate groups (**32**), the yields of the bromination and azide substitution reactions that yield azide-containing **33** were slightly improved (Scheme 19c, bottom). This route required several more alcohol deprotection and protection steps (**34**-**36**) before yielding the final target (**29**). However, despite the increase in reaction steps, the overall yield improved significantly and was more scalable than the previous route.

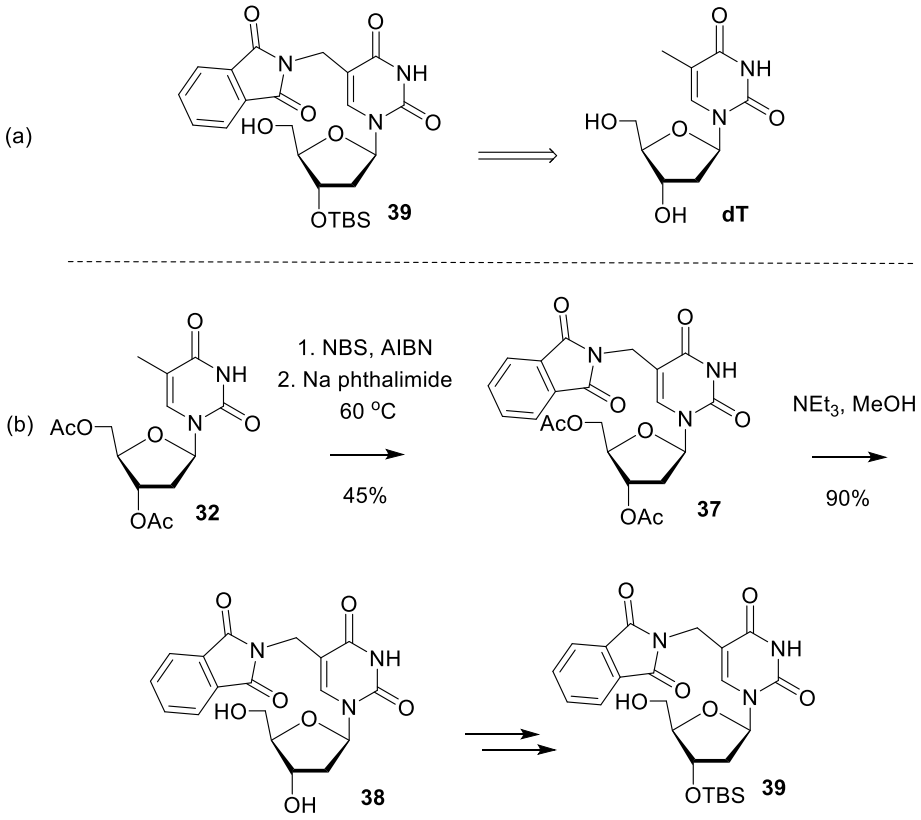
Scheme 19. Retrosynthetic (a) and forward (b and c) synthesis of azide-containing nucleobase via 2'-deoxyuridine (b) and thymidine (c).



Unfortunately, reduction of the C5-azide, necessary for functionalization, using triphenyl phosphine (PPh_3) was unsuccessful. No reaction was observed when the azide (i.e. **29**, **33**, or **36**) was reacted with PPh_3 under multiple conditions (i.e. solid PPh_3 or PPh_3 on a polymer support). While hydrogenation did reduce the azide in the monomers (i.e. **29**, **33**, **36**), this reaction was

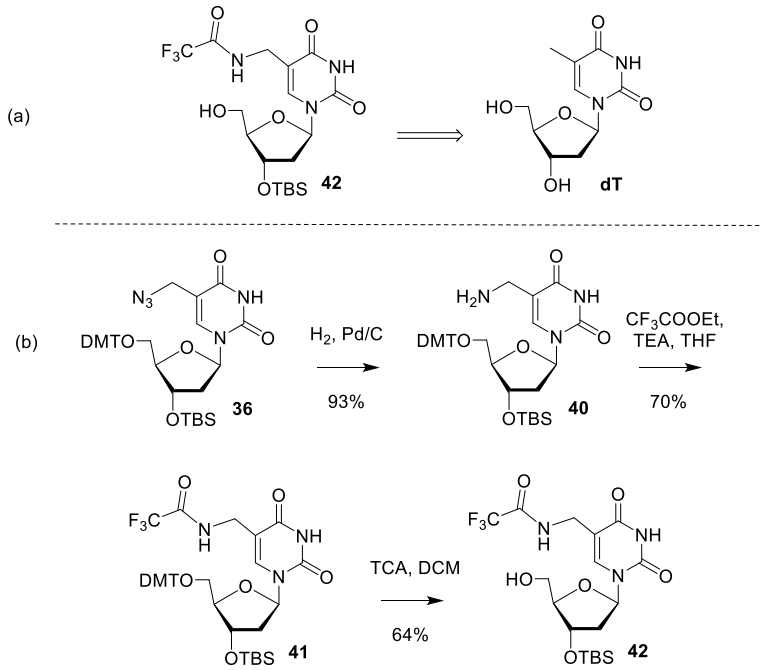
incompatible with the pentenoyl-protected DOB structure, because the pentenyl group(s) is susceptible to reduction by H_2. Furthermore, presumably due to the same Staudinger reaction, phosphoramidite coupling in the presence of an azide resulted in a very low yield (i.e. 20%) because P(III) was easily oxidized by N_3. To avoid the azide, we tried different ways to introduce a protected amine, including as a phthalimide or trifluoroacetamide. The phthalimide was introduced in a similar manner as the azide; utilizing sodium phthalimide to displace the bromide to yield **37** (Scheme 20). The acetate protecting groups were efficiently deprotected to yield **38**. As a model deprotection, the phthalimide group in either **37** or **38** to reveal the primary amine was attempted under multiple conditions (e.g. H_2NNH_2, $MeNH_2$), but did not yield the desired product based on mass spectrometry and NMR. The identity of the reaction product(s) was never determined, and the synthesis of **39** was abandoned.

Scheme 20. Retrosynthetic (a) and forward (b) synthesis of a phthalimide containing nucleobase.



After introducing the protected amine as the phthalimide was unsuccessful, we turned to a trifluoroacetamide group to mask the amine. The trifluoroacetamide protecting group was introduced by reduction (**40**), protection (**41**), and detritylation (**42**) of previously synthesized **36** (Scheme 21).

Scheme 21. Retrosynthetic (a) and forward (b) synthesis of a trifluoroacetamide containing nucleobase.

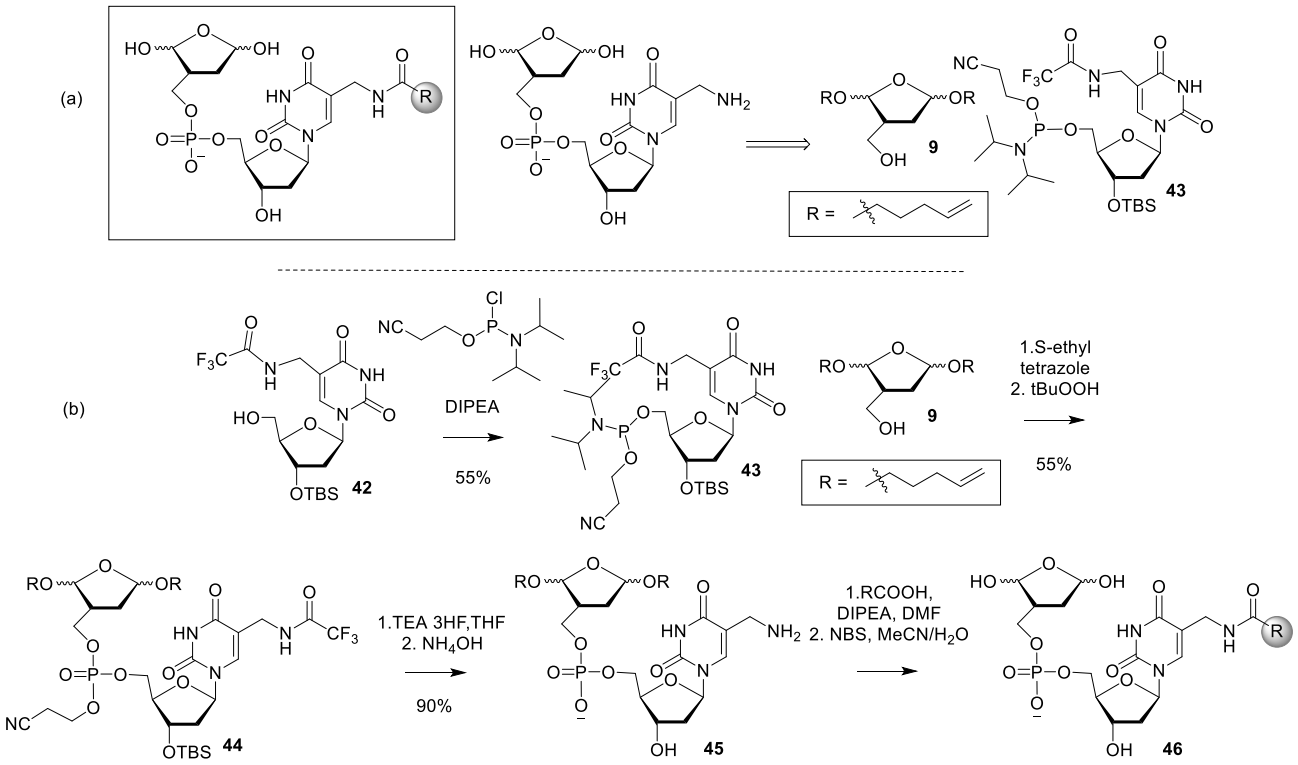


The first-generation library was synthesized using this monomer (Scheme 22). Compound **42** was phosphitylated (**43**) and coupled to previously synthesized **9** (Scheme 12) to yield **44**. Following deprotections, the scaffold molecule (**45**) was coupled to a variety of carboxylic acids (140 compounds) and the pentenoyl groups were removed with NBS to generate the library of inhibitor candidates (**46**).

Overall, the synthesis of the first-generation library bearing structural diversity at the 5-methyl position of the nucleobase is twelve steps with an overall yield of 10%. The limiting step in this synthesis remains the bromination reaction required for incorporating a functional group at

the 5-methyl position. The bromination was low yielding and was not scalable above 1 g. To overcome this hurdle, multiple bromination reactions were carried out in parallel and combined so the final yield did not suffer due to one step. Altogether, approximately 150 mg of the final scaffold (**45**) was synthesized and a portion (i.e. 40 mg) was used to make a library.

Scheme 22. Retrosynthetic (a) and forward (b) synthesis of amine-containing scaffold (**45**).



3.2 In vitro biochemical assays

Upon library generation, biochemical assays were used to evaluate the inhibitory activity of the individual compounds. There were multiple ways to do so including qualitative high throughput assays and more laborious quantitative assays. In addition, Pol β is a bifunctional enzyme and its polymerase and lyase activities were analyzed separately. This is important because the inactivation of one activity does not necessarily translate to inactivation of the other. In other

words, the inactivation of the enzyme's lyase and polymerase activity are not mutually inclusive. Furthermore, we could not predict which activity or activities the candidate inhibitor(s) would inhibit.

3.2.1 Fluorescence-based strand displacement assay

This high-throughput screening assay indirectly analyzed Pol β's polymerase activity via a strand displacement assay (Figure 24).[158] This assay utilized a ternary oligonucleotide complex containing a 3'-fluorophore (i.e. carboxytetramethylrhodamine (TAMRA)) and a 5'-quencher (i.e. Black Hole Quencher (BHQ), Figure 24). This assay was indirect because rather than detecting the incorporation of a nucleotide triphosphate, it measured the increase in fluorescence signal when the strand containing a fluorophore was displaced from the strand containing a quencher because of polymerization. As Pol β incorporated dNTPs in the 1 nt gapped DNA substrate (**B**, Table 1, Section 3.2.4), the strand containing the fluorophore was displaced, leading to an increase in fluorescence. Therefore, inhibitors were identified by a decrease or absence of fluorescence increase.

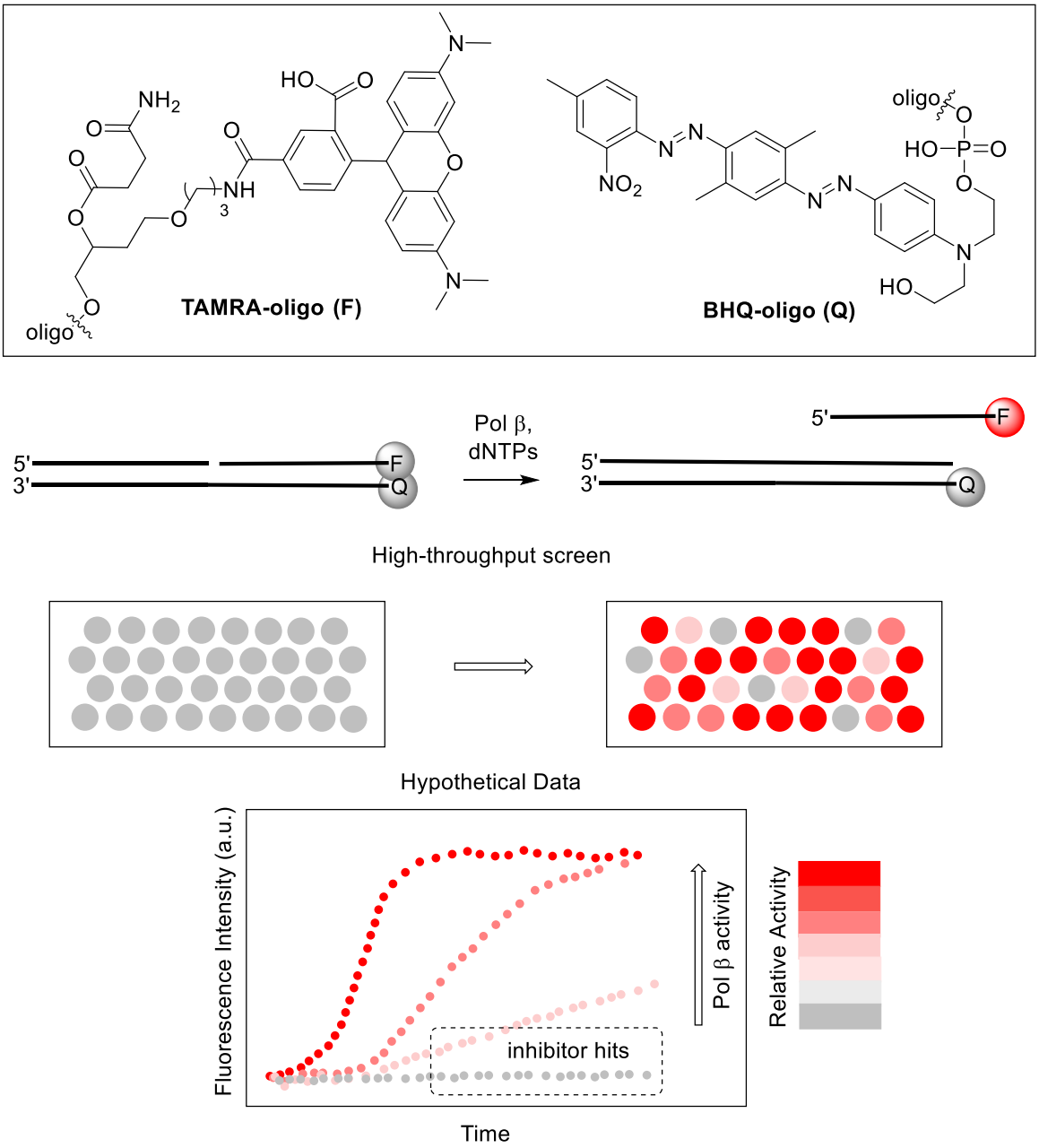


Figure 24. Fluorescence-based strand displacement assay.

This assay was originally used as a qualitative method to identify potential hits that must be analyzed further using more quantitative biochemical assays. Later, the assay was adapted to extract kinetic information from polymerase inhibitors. The data were fit to a single exponential growth equation (1) that follows a plateau. The plateau was important because the strand

displacement assay exhibited an induction period in which several nucleotides of the fluorescently labelled DNA were displaced before the fluorescently labeled oligonucleotide was released into solution. This induction time was determined by inspection and typically varied between 10 and 15 min. The data were fit beginning at the time when a growth in fluorescence was observed.

$$Y = F_0 + (F_1 - F_0) \times (1\text{-}e^{-kt}) \quad (1)$$

Y is the fluorescence intensity, F_0 is the fluorescence value at time 0, F_1 is the fluorescence value at time ∞, k is the rate constant, and t is time. Rate constants are extracted for each experiment and relative rates are determined using equation (2).

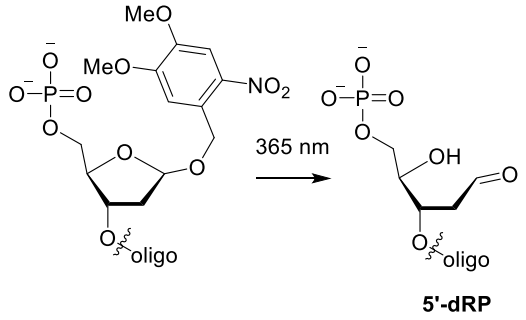
$$k_{rel} = k_{inhibitor}/k_{pol\beta} \quad (2)$$

Where $k_{inhibitor}$ is the rate constant for experiments containing inhibitor and $k_{pol\beta}$ is the rate constant for control experiments lacking inhibitor.

3.2.2 Gel-based lyase assay

The gel-based lyase assay directly monitored Pol β excision of the 5'-dRP group in a ternary DNA complex (Figure 25). This assay utilized 3'-^{32}P labeled DNA containing a 5'-dRP that was generated from an oligonucleotide containing a photolabile precursor to this unstable molecule (Scheme 23).[29] An oligonucleotide with the precursor was radiolabeled and hybridized with template and flanking strands to form a ternary complex (**A**, Table 1, Section 3.2.4). The Nv protecting group was cleaved photolytically to reveal the 5'-dRP (Scheme 23).

Scheme 23. Cleavage of Nv to obtain 5'-dRP.



This assay relied on ^{32}P-radiolabeling and was more sensitive than the previously described fluorescence-based assay. Compared to the strand displacement assay, a smaller amount of DNA was needed to monitor the reaction. In the presence of active Pol β, the 5'-dRP group was efficiently removed and the loss was visualized by gel electrophoresis in the form of a faster moving product (Scheme 7). The linear relationship between the amount of product versus time was used to determine the reaction rate. The relative enzyme activity was determined using equation 2 (Section 3.2.1). In the case of inactive Pol β, the amount of product does not increase as a function of time (gray data, Figure 25).

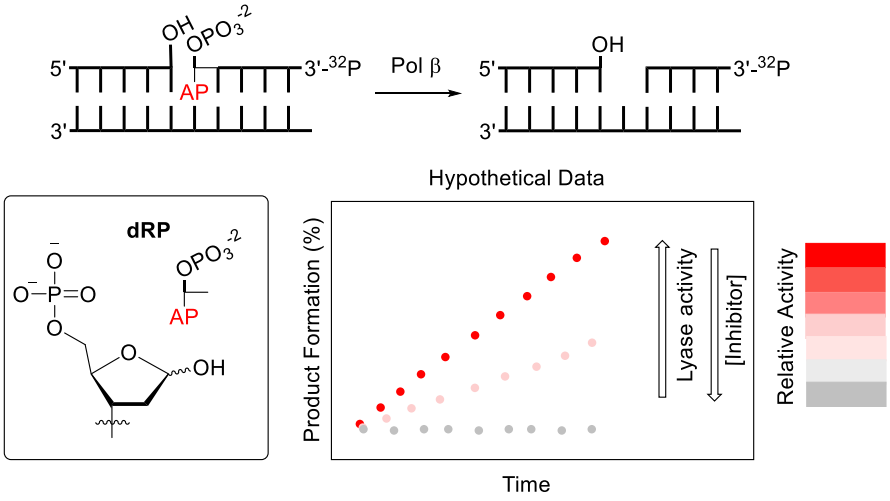


Figure 25. Gel-based lyase assay.

Relative rates are essential for extracting kinetic information and quantitative analysis. There are three major kinetic parameters used to describe and quantify the efficiency of an inhibitor. The first being an IC_{50}, which represents the concentration of inhibitor that results in 50% enzyme activity and is more often utilized to characterize reversible inhibitors. For an irreversible inhibitor, an IC_{50} plot used relative enzyme activity across multiple inhibitor concentrations at a single preincubation time (Section 3.3.2). Because inhibition by an irreversible inhibitor is dependent on preincubation time, an IC_{50} of a covalent inhibitor varies as this variable

is changed. The concentration of inactivator at half-maximal inactivation (K_I) and the maximal rate constant for inactivation (k_{inact}) are two other quantitative parameters that better describe an irreversible inhibitor. To obtain these values, the effect of inhibitor concentration and preincubation time on reaction velocity must be explored. This analysis required several kinetic experiments (Figure 26).

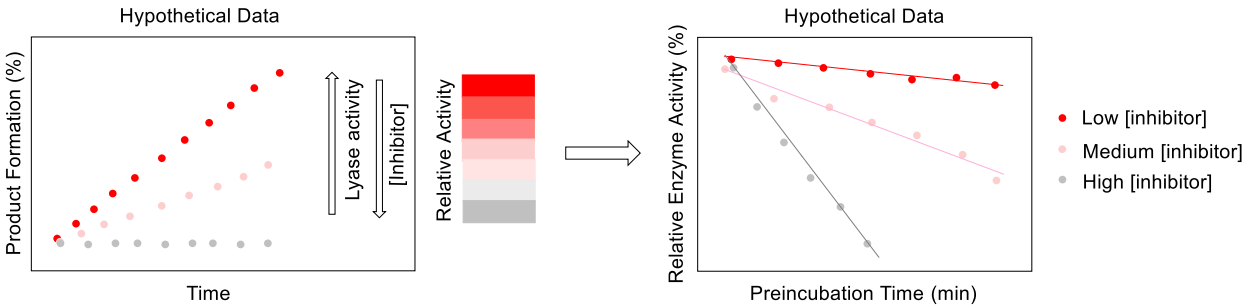


Figure 26. Transforming hypothetical raw data (left) into a preincubation plot (right).

Each data point in a preincubation plot comes from one set of experimental conditions in which product formation was measured over several time points, completed in triplicate. In the case of an irreversible inhibitor, the relationship between relative enzyme activity and preincubation time was linear at a single inhibitor concentration (equation 3).

$$A = mt + b \quad (3)$$

A Kitz-Wilson plot was valuable for evaluating covalent inhibition.[159] A Kitz-Wilson plot showed the linear relationship between the half-life ($t_{1/2}$ from multiple concentrations) and the inverse of the inhibitor concentration. The half-life of the enzyme ($t_{1/2}$) was determined by using the y-intercept (b) and the slope of the linear relationship (m) to solve for the preincubation time (t) at which 50 % enzyme activity (A) remained (equation 3), based on the preincubation plot

(Figure 26). The x- and y-intercepts of a Kitz-Wilson plot were used to calculate the K_I and k_{inact} using equations 4 and 5, respectively.

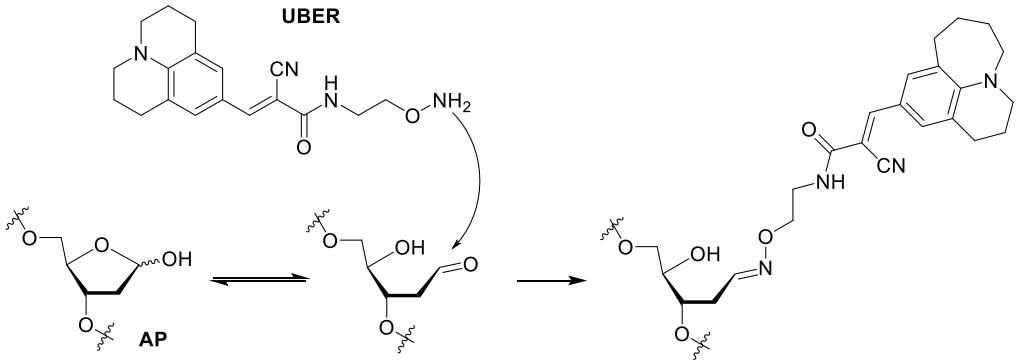
$$\text{x-int} = -1/K_I \quad (4)$$

$$\text{y-int} = \ln(2)/k_{inact} \quad (5)$$

3.2.3 Fluorescence-based lyase assay

Initially, the previous two assays were our primary biochemical assays for analyzing Pol β activity. In 2019, a novel fluorescence-based probe was reported for detecting BER glycosylase activity.[160] The universal BER (UBER) probe contains a fluorophore connected to a reactive alkoxyamine. The alkoxyamine reacts with aldehydes. This assay was originally designed to detect an abasic site (AP) that forms after an *N*-glycosylase removes a damaged base, such as dU (Figure 13). The probe is rigidified upon reaction with the aldehyde in the context of duplex DNA (Scheme 24) and the loss of free movement (bond rotation) yields an increase in fluorescence signal (Figure 27).

Scheme 24. Reaction between AP and UBER Probe.



Considering that the chemistry of the UBER probe relied on reaction between the alkoxyamine and a reactive aldehyde, we hypothesized that this assay could be adapted to detect 5'-dRP lyase activity in a high-throughput manner (Figure 27). However, we first had to determine if the UBER assay would work for detecting 5'-dRP. We were unsure whether the greater flexibility of an incised DNA containing 5'-dRP would give rise to too small of an increase in fluorescence upon reaction with the UBER probe, reducing the sensitivity of the assay.

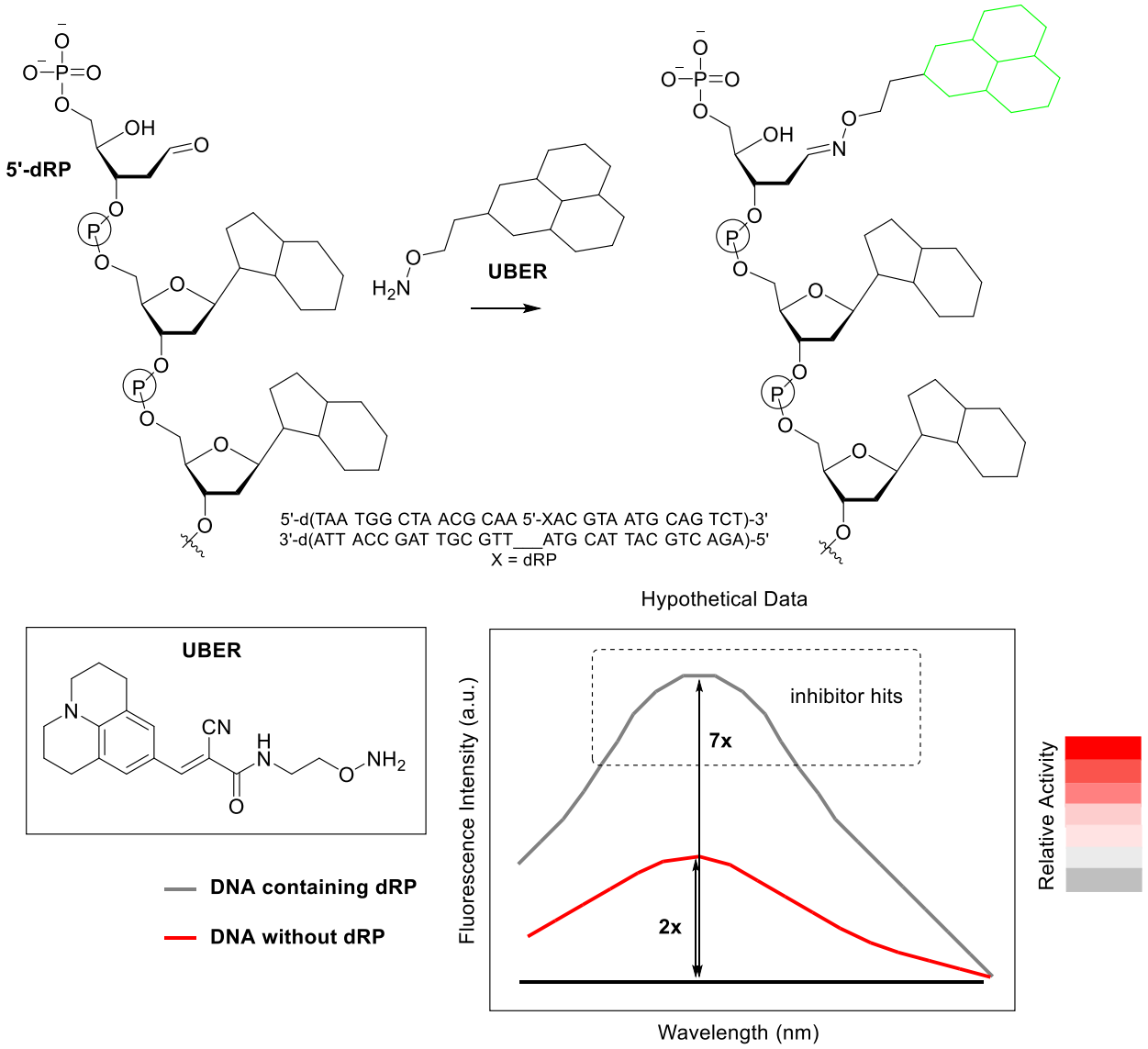


Figure 27. Fluorescence-based UBER lyase assay.

To adapt this assay for our inhibitor screen, the UBER probe was reacted with a ternary complex containing 5'-dRP (Figure 27). To make this assay viable for high-throughput screening, we considered the following:

1. The lowest DNA concentration needed to detect a signal is determined by the sensitivity of the fluorescence measurement.
2. The ratio of DNA to Pol β must be optimized. A small ratio contributes to efficient processing of DNA by active Pol β. However, a large ratio requires less enzyme. It is desirable to use a low concentration of enzyme during the screen because a high concentration of enzyme requires a higher concentration of inhibitor for inactivation (Section 3.3). Higher inhibitor concentrations are less desirable when trying to identify a good hit compound.
3. The ratio of DNA to probe must be optimized. The DNA must be limiting relative to the probe to ensure the probe reaction goes to completion. However, we were unsure if a large excess of probe may produce a high background signal. Therefore, the background fluorescence signal of the probe must be minimal.

Altogether, after considering the sensitivity of the probe with 5'-dRP and the concentration ratios of the different components (i.e. Pol β < DNA < UBER probe), the UBER probe was reacted with a ternary complex containing 5'-dRP (**A**, Table 1, Section 3.2.4). The background signal of the UBER probe at 5 μM was minimal (left black bar, Figure 28). The fluorescence response of the probe was evaluated with a range of DNA concentrations with or without reaction with Pol β. Several enzyme quenching methods, such as pH, EDTA, and SDS, were examined (Figure 28). We determined that the background signal of the UBER probe alone was low. A 5:1 concentration

ratio between the UBER probe (5 μM) and 5'-dRP DNA (1 μM) was enough to elicit a detectable fluorescence change. Furthermore, in the presence of Pol β (100 nM), the fluorescence signal decreased 2.5-fold.

The lyase activity of Pol β decreased the amount of dRP, reducing the fluorescence signal. In contrast, an inhibitor was identified by a high fluorescence signal. This method was qualitative and identified potential inhibitors but required quantitative analysis after the initial screen (Section 3.2.2). Lastly, we found that any quenching method we tried effectively quenched the enzyme, so we chose pH 6.5 for our screen because it simultaneously quenched the enzyme and slightly improved the efficiency of the UBER reaction (Figure 28).

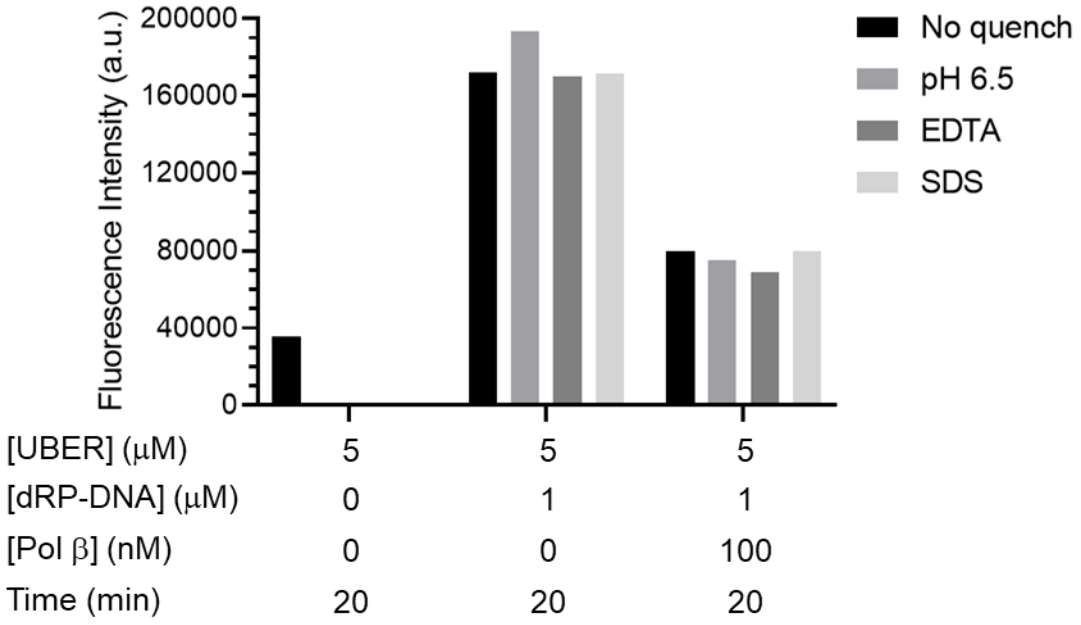


Figure 28. Conditions to detect dRP using UBER assay. DNA **A** containing dRP (5 μM) was incubated with or without Pol β (100 nM) in 1X reaction buffer (50 mM HEPES, 50 mM KCl, 5 mM $MgCl_2$, 200 mM EDTA, 0.01% Tween 20, pH 7.5) for 20 min. After preincubation time, the solution was diluted 5-fold, quenched with various conditions (pH 6.5, 1% SDS or 10 mM EDTA) and mixed with the UBER probe (5 μM) for 4-5 h. After incubation, fluorescence was measured using a Clariostar well plate reader.

3.2.4 Assays and DNA substrates used for different polymerases

The previous assays described were most commonly used when screening and quantifying library molecules against Pol β. However, different assays were used for *in vitro* selectivity assays because different polymerases have different substrate preferences and activities. The other polymerases used for determining selectivity were Klenow exo$^-$, Pol θ, Pol η, and Pol λ.

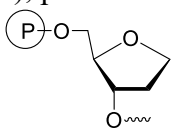
These polymerases were chosen because their structures and activities were diverse compared to Pol β. Klenow exo$^-$ is a model replicative polymerase, while the remaining have roles in mammalian DNA repair. The selectivity for Pol β and Pol λ was especially of interest because Pol λ is the back-up polymerase in BER and the two have a high sequence homology (Section 2.3.1).[98–100]

The DNA substrate for each assay depended on the activity and preferences of each enzyme (Table 1). Pol β preferred nicked ternary complexes (**A** and **B**, Table 1). In order to monitor lyase activity, the nick bore a 5'-dRP (**A**). Substrate **B** was also used for Klenow exo$^-$ because the assay was quick and convenient. Primer extension assays were used to test Pol θ and Pol η. However, the primer (**C**) for Pol θ was a single strand of DNA and SYBR Gold was used to measure an increase in fluorescence that corresponds to newly synthesized DNA.[161] The primer (**D**) for Pol η was part of a duplex and was 5'-^{32}P-labeled to detect nucleotide incorporation.[162,163] Like Pol β, Pol λ preferred nicked and gapped ternary complexes. Pol λ did not have sufficient lyase activity to quantify using substrate **A** so we used a 2-nt gapped complex (**E**) and measured the incorporation of two nucleotides.[25]

An additional substrate (**F**) was used for fluorescence anisotropy experiments.[135] This substrate was not used for selectivity but rather to determine DNA binding by Pol β and its distinct domains.

Table 1. DNA substrates for assays.

Pol β[1] (**A**)	5'- d(TAA TGG CTA ACG CAA XAC GTA ATG CAG TCT) -3' 3'- d(ATT ACC GAT TGC GTT__ATG CAT TAC GTC AGA) -5'
Pol β[2] and Klenow exo$^-$ (**B**)	5'- d(TCA CCC TCG TAC GAC TC TTT TTT TTT TGC F) - 3' 3'- d(AGT GGG AGC ATG CTG AG__AAA AAA AAA ACG Q) - 5'
Pol θ (**C**)	5'- d(TTT TTT TAG GTT T) -3'
Pol η (**D**)	5'- d(TCA CCC TCG TAC GAC TC) -3' 3'- d(AGT GGG AGC ATG GTC AGG ACC T) -3'
Pol λ (**E**)	5'- d(ACC ATG GGA CGT GCT G ACT CCA CTA GAT ACA CTT) -3' 3'- d(TGG TAC CCT GCA CGA CTC TGA GGT GAT CTA TGT GGA) -5'
Fluorescence Anisotropy (**F**)	5'- d(TAA TGG CTA ACG CTT pFCC GTA ATG CAG TCT) -3' 3'- d(ATT ACC GAT TGC GAA__AGG CAT TAC GTC AGA Fl) -5'

X = dRP, F = TAMRA, Q = BHQ, Fl = dichloro-diphenyl-fluorescein (SIMA-HEX), pF = 
[1] = lyase assay (Section 3.2.2 and Section 3.2.3), [2] = fluorescence assay (Section 3.2.1)

3.3 First-generation Pol β inhibitor(s) and their domain selectivity

The first goal of this research was to identify a potent and selective inhibitor for WT Pol β. As mentioned before, previous inhibitors exhibit several shortcomings (Section 2.6). Therefore, we aimed to identify and fully characterize inhibitors using kinetics, mass spectrometry and binding assays. First-generation inhibitors represent a starting point but we anticipated that they would not be potent enough to warrant extremely detailed characterization. However, they are a necessary step in identifying a second-generation inhibitor (Figure 23).

An important aspect of screening and evaluating inhibitors lies in the preincubation mixture. Several components can affect the inhibitory activity of a compound, including:

- The concentration of enzyme in the preincubation and reaction mixture

- The ratio of inhibitor to enzyme
- The amount of dilution from preincubation mixture to reaction mixture
- Preincubation time

As mentioned before, a low concentration of enzyme during preincubation was desirable because it required less inactivator to show an inhibitory effect (Section 3.2.3). The concentration of enzyme during preincubation was also determined by the reaction assay. The lower limit of enzyme concentration during the reaction mixture was determined by the need for there to be enough enzyme to measure activity. This limit was defined by the sensitivity of the assay. We determined that 5 nM Pol β was sufficient in our assays to monitor lyase and strand displacement activity.

The ratio of inhibitor to enzyme was important because a larger ratio yields a greater inhibitory effect. However, it was impractical to use an extremely large ratio for the purposes of identifying potent inhibitors. An inhibitor that inactivated an enzyme at 5 mM was not very impressive. We were interested in compounds that inactivate Pol β at low μM or, even better, nM concentrations. Therefore, the ratio of inhibitor to enzyme was optimized during a screen to yield a small, manageable number (i.e. < 10) of potent hit compounds.

The preincubation dilution was also important when measuring enzyme activity. A large preincubation dilution was desirable because the rate at which the inhibitor inactivated the enzyme should be as low as possible while monitoring the enzyme catalyzed reaction. Furthermore, if the inhibitor concentration was high enough to continuously inactivate the enzyme during kinetic measurements, the concentration of active enzyme would continue to decrease during the assay. The limitations of the preincubation dilution were determined by the amounts of enzyme and inhibitor during the preincubation and reaction mixtures (described previously). Enzyme kinetics

were carried out at 5 nM Pol β. We determined that a 50 nM Pol β preincubation mixture and a 10-fold dilution generally satisfied all considerations. It should be noted that in some cases, the amount of dilution and concentration of enzyme was adjusted to fit the specifications of a certain assay.

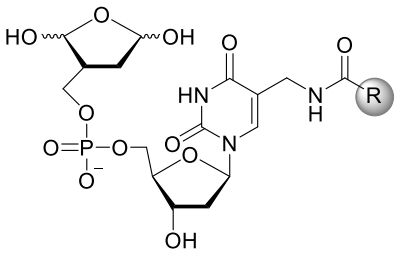
As mentioned previously, irreversible inhibition depends on the amount of time spent preincubating with the enzyme (Section 3.2.2). Therefore, preincubation time must be long enough to yield inhibition but short enough to still be convenient and easy to work with. Generally, we used a 30 min preincubation time during the screening process because handling so many library compounds (~80 at a time) required roughly 4-5 min to properly mix reagents (e.g. Pol β, inhibitor, buffers, etc.) so a longer preincubation prevented experimental error. Once a hit was identified, we found that 20-30 min preincubations were sufficient for our experiments. At shorter preincubation times (e.g 5-10 min), we observed less inhibition at optimal inhibitor concentrations.

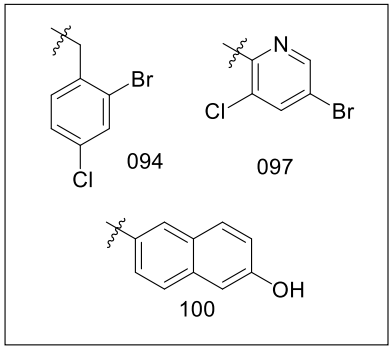
3.3.1 Lyase domain selectivity

Amine scaffold **45** (total 40 mg, 200 nmol per carboxylic acid coupled, Scheme 22) was functionalized with 280 different carboxylic acids, deprotected, and analyzed via the strand displacement assay (Figure 24). The amide bond reaction was optimized using standard activating agents (1.4 equiv. HBTU, 1.4 equiv. HOBt), base (20% v:v DIPEA), and a slight excess of the carboxylic acid (1.4 equiv.). These reactions were most effective when the amine (**45**) and the carboxylic acid were azeotropically dried together prior to adding the other reagents. Following amide coupling, the DOB component was unmasked using NBS (2.5 equiv.) in 30% H_2O in acetonitrile. These library molecules were screened against Pol β using the fluorescence-based strand displacement assay (Section 3.2.1). Each compound (25 μM) was preincubated with Pol β (100 nM) for 30 min, mixed with a 5X cocktail solution containing dTTP and fluorescently

labeled ternary complex (Figure 24; **B**, Table 1), and diluted further (10X for preincubation mixture, 5X for DNA) in 1X buffer in the assay reaction. The initial screen often identified several inhibitors, so all hit compounds were resynthesized and rescreened multiple times. False positives were ruled out by comparing inhibition induced by the hit compounds with inhibition induced by the free carboxylic acid(s) in the presence of all activating agents but no scaffold (**45**).

This library (**46**) led to the identification of two hits (**47** and **48**, Figure 29). Two other compounds (094 and 097, Appendix 1) were identified in this screen, however, after additional assays, these were determined to be false positives. Compounds **47** and **48** were by far the most potent hits identified in the library (Appendix 1). Compounds **47** and **48** are regioisomers of each other. The structure of the hits identified was intriguing because the hydroxylated naphthoic acid moiety is present in other Pol β inhibitors (i.e. **4**, Figure 22 and PA, Figure 20).[27,145] We compared these two isomers with a library compound (100, Appendix 1) that was not identified in the screen, despite their similar structures (hydroxylated naphthoic acid). However, it was clear that the library compound that arose from coupling with 6-hydroxyl-2-naphthoic acid (100, Appendix 1) was not a hit. This suggested that the 3-hydroxyl group of the naphthoic acid may be critical for binding, considering this group is present in the inhibitors identified, as well as **4** and PA.





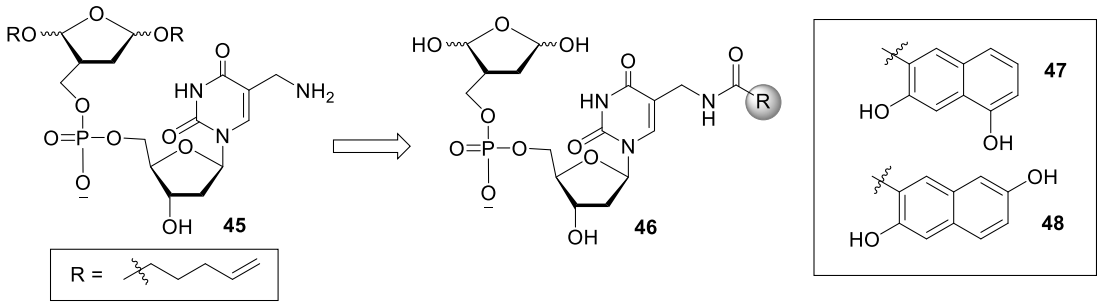
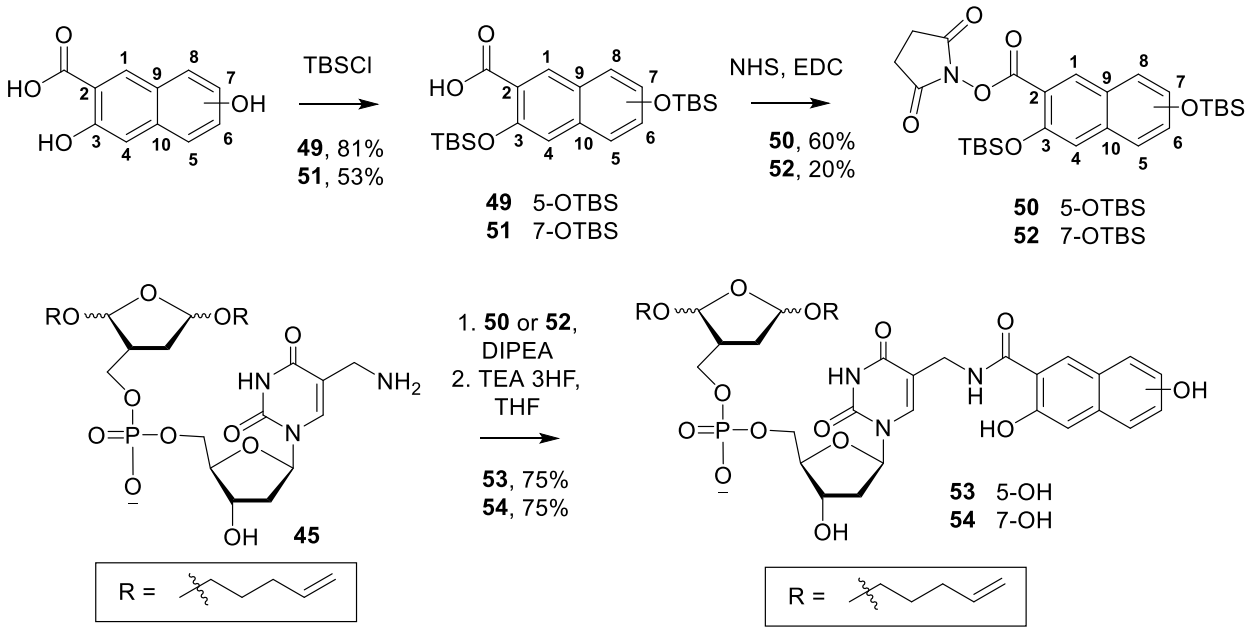


Figure 29. First-generation inhibitor hit compounds **47** and **48**.

The library screen was completed with crude material. In order to analyze hit compounds quantitatively, **47** and **48** were synthesized on a larger scale and the pentenoyl protected analogues were purified. The amide coupling was improved by silylating the carboxylic acids (**49** and **51**) and coupling them as their NHS esters (**50** and **52**) (Scheme 25). Coupling scaffold **45** and the NHS ester (**50** or **52**) and subsequent desilylation yielded the pentenoyl protected precursors (**53** and **54**) of the active inhibitor hits (**47** and **48**).



Scheme 25. Improved synthesis of the hit compounds.

Following a 30 min preincubation, compounds **47** and **48** completely inactivated Pol β's strand displacement activity at around ~ 5 and 10 μM respectively (Figures 30 and 31). The curves associated with concentrations ≥ 8 μM are hidden behind the control associated with no enzyme (blue, Figures 30-31). Strand displacement assays were used as a qualitative tool to measure polymerase activity. We turned to the gel-based lyase assay to monitor the enzyme's lyase activity (Figure 25, Section 3.2.2).

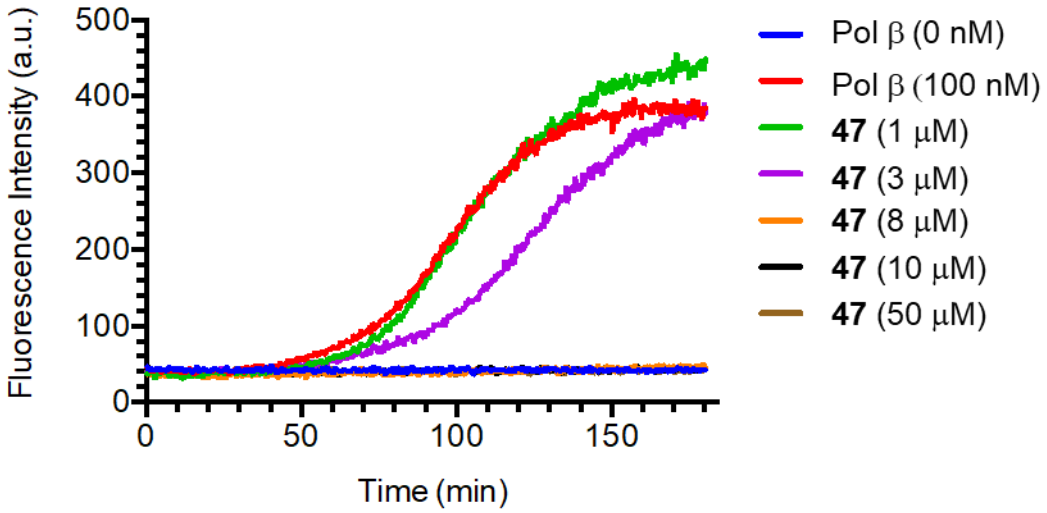


Figure 30. Strand displacement assay of Pol β with **47**. Pol β (100 nM) was preincubated with hit **47** at various concentrations (listed) for 30 min. The preincubation mixture was diluted 10-fold and mixed with DNA cocktail solution, to yield 5 nM Pol β, 50 nM DNA substrate **B,** and 100 μM dTTP**.**

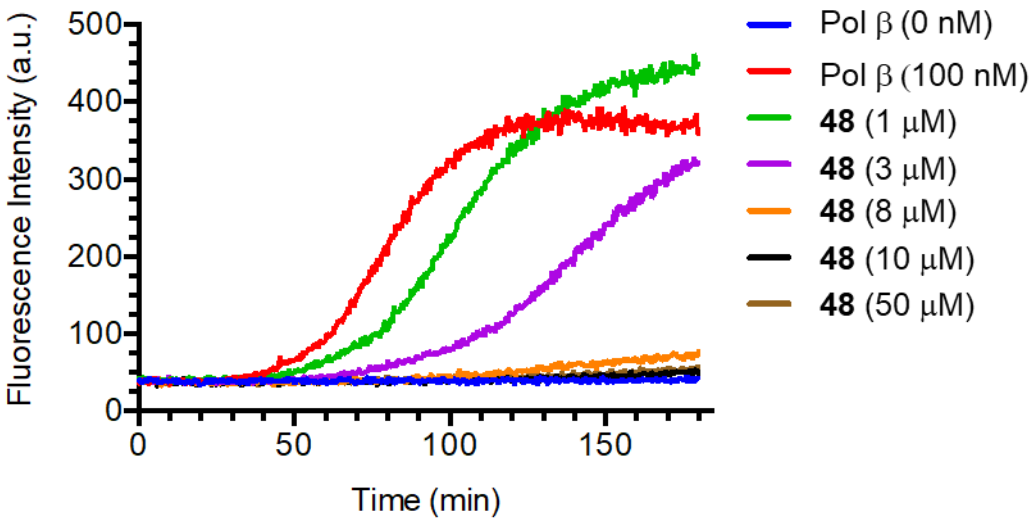


Figure 31. Strand displacement assay of Pol β with **48.** Pol β (100 nM) was preincubated with hit **48** at various concentrations (listed) for 30 min. The preincubation mixture was diluted 10-fold and mixed with DNA cocktail solution, to yield 5 nM Pol β, 50 nM DNA substrate **B,** and 100 μM dTTP.

Inhibitors **47** and **48** inactivate the 8 kDa lyase domain at 25 and 50 μM, respectively (Figure 32). A larger inhibitor concentration was used for these assays, in comparison to the strand displacement assays (Figures 30-31) because the 8 kDa domain was slightly less active than the WT Pol β. To accommodate the decrease in activity, the lyase activity was measured at a higher concentration of 8 kDa Pol β and the inhibitor was scaled accordingly.

Lyase inhibition was determined by using a single preincubation time (30 min) and time point (15 min) (Figure 32). These data are still qualitative and so to attain quantitative data, we measured the lyase activity as a function of time with multiple inhibitor concentrations and preincubation times (Figures 33-35). As expected, the relative activity decreased in response to longer preincubation times across all inhibitor concentrations of **47** and **48.** As mentioned before, this behavior was indicative of irreversible inhibition (Section 3.2.2). Furthermore, this result was evident in both the intact enzyme (**47,** Figure 33; **48**, Figure 35) and the 8 kDa domain (**47,** Figure 34). Both the strand displacement and lyase assays indicated that **47** was the more potent inhibitor.

Each preincubation plot (Figures 33-35) required at least one week of experiments (Section 3.2.2). Considering inhibitor **48** was not as potent as **47** in the holo-enzyme, we chose not to evaluate the lyase activity of the 8 kDa domain as a function of time with multiple preincubation times and concentrations of **48**.

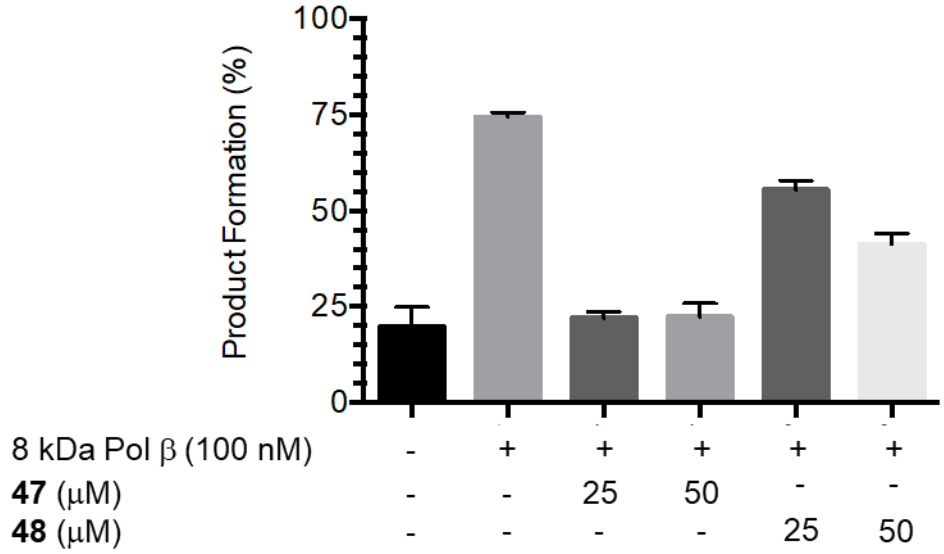


Figure 32. Lyase activity in the presence of inhibitor hits **47** and **48**. 8 kDa Pol β (100 nM) was preincubated with BSA (0.01 mg/mL) and either **47** or **48** at various concentrations (listed) for 30 min. Preincubation mixture was diluted 10-fold and mixed with 25 nM DNA substrate **A**. Reaction was completed in triplicate, stopped after 15 min of reaction time, quenched with 150 mM $NaBH_4$, and analyzed by gel electrophoresis to quantify product formation of the lyase reaction.

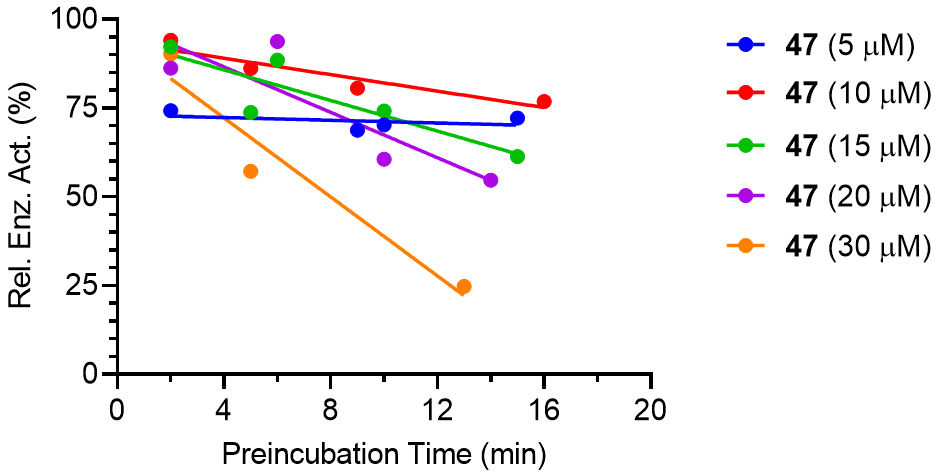


Figure 33. Irreversible lyase inhibition of WT Pol β lyase activity by **47.** Pol β (50 nM) was preincubated with **47** at various concentrations (5, 10, 15, 20, 30 μM) for a range of time. Each point represents a reaction of certain inhibitor concentration and preincubation time. The preincubation mixture was diluted 10-fold and mixed with 50 nM DNA substrate **A**, stopped at various time points, quenched with 150 mM $NaBH_4$, and analyzed by gel electrophoresis to quantify product formation of the lyase reaction. Relative rates of 50 nM Pol β in the presence of **47** at various concentrations (listed) compared to the rate of 50 nM Pol β in the absence of inhibitor.

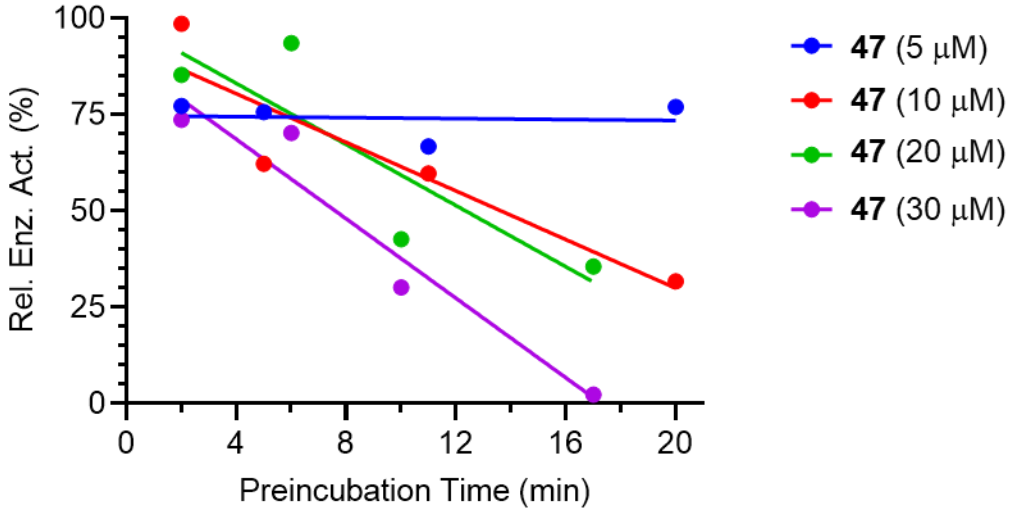


Figure 34. Irreversible lyase inhibition of WT Pol β 8 kDa domain lyase activity by **47.** 8 kDa Pol β (100 nM) was preincubated with **47** at various concentrations (5, 10, 20, 30 μM) for a range of time. Each point represents a reaction of certain inhibitor concentration and preincubation time. The preincubation mixture was diluted 10-fold and mixed with 25 nM DNA substrate **A**, stopped at various time points, quenched with 150 mM $NaBH_4$, and analyzed by gel electrophoresis to quantify product formation of the lyase reaction Relative rates of 100 nM Pol β 8 kDa domain in the presence of **47** at various concentrations (listed) compared to the rate of 100 nM Pol β 8 kDa domain in the absence of inhibitor.

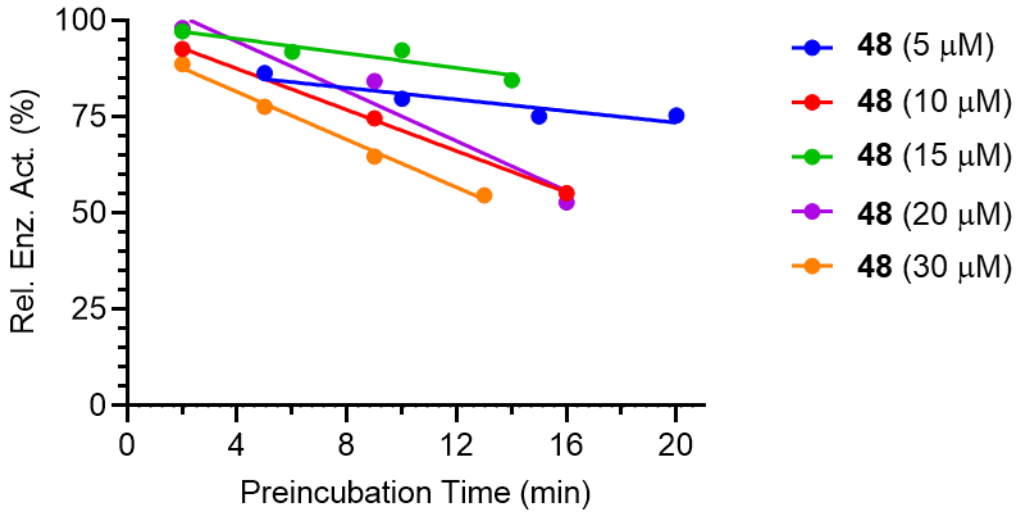


Figure 35. Irreversible lyase inhibition of WT Pol β lyase activity by **48.** Pol β (50 nM) was preincubated with **48** at various concentrations (5, 10, 15, 20, 30 μM) for a range of time. The preincubation mixture was diluted 10-fold and mixed with 50 nM DNA substrate **A**, stopped at various time points, quenched with 150 mM $NaBH_4$, and analyzed by gel electrophoresis to quantify product formation of the lyase reaction. Relative rates of 50 nM Pol β in the presence of **48** at various concentrations (listed) were compared to the rate of 50 nM Pol β in the absence of inhibitor.

While the preincubation plots (Figures 33-35) suggested that the inhibitors act in an irreversible manner, the data cannot be used to extract quantitative information, like IC_{50}, $K_{I,}$ or k_{inact}. The problem lied in the relative enzyme activity across multiple inhibitor concentrations. As the concentration of the inhibitor increased, the linear relationship between relative enzyme activity and preincubation time was more pronounced (i.e. steeper slope). However, when looking at a single preincubation time, the relative activity associated with each inhibitor concentration was not what was expected. For example, a lower concentration of inhibitor was expected to have the highest relative enzyme activity across all preincubation times when compared to experiments in which a higher inhibitor concentration was used. However, this is clearly not the case with **47** and **48**. As a result, any IC_{50} or Kitz-Wilson plot extrapolated from these data was nonsensical.

Therefore, we re-evaluated lyase inactivation by **47** using a 30 min preincubation time across multiple inhibitor concentrations. The elimination of multiple preincubation times significantly reduced the time required to complete this experiment. By measuring relative activity at a single time point (15 min), we determined **47** inhibited WT Pol β with and IC_{50} value of 18.7 μM (Figure 36). This is a slight improvement from previous first-generation inhibitor **2**, which had an IC_{50} value of 21 μM against WT Pol β (Figure 21).[26] Unsurprisingly, neither first-generation inhibitor (i.e. **2**, **47**) exhibited similar potency to second-generation Pol β inhibitor **4** (IC_{50} = 400 nM, Figure 22).[27] However, it is worth noting that the IC_{50} value of inhibitor **4** was determined using different conditions (i.e. smaller preincubation dilution).[27]

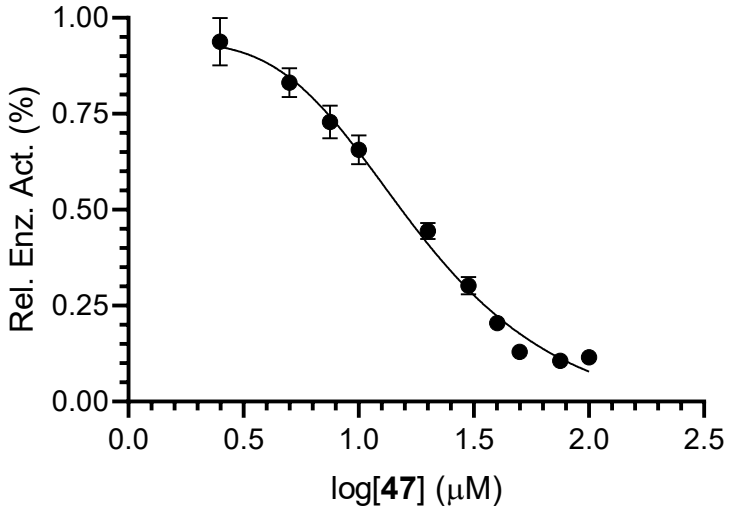


Figure 36. Inhibitor **47** has an IC_{50} value of 18.7 μM in WT Pol β.

The irreversible inactivation by **47** was confirmed by dialysis experiments with Pol β (Figure 37). Untreated Pol β retained 70% activity after dialysis but, in the presence of **47** (80 μM), Pol β was completely inactivated and did not regain any activity after dialysis.

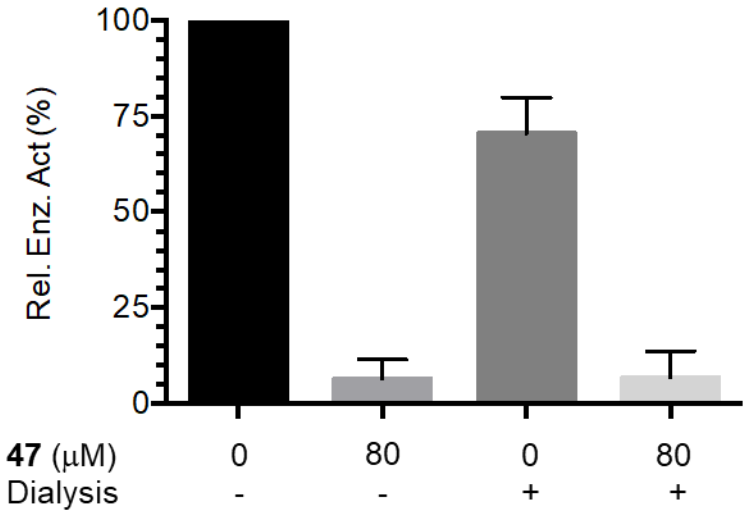


Figure 37. Lyase activity of WT Pol β with **47** before and after dialysis. Pol β (100 nM) was preincubated with or without **47** (80 μM) for 30 min. An aliquot of preincubation mixture (2.5 μL) was diluted 10-fold and mixed with 150 nM DNA substrate **A**. Reaction was completed in triplicate, quenched with 150 mM $NaBH_4$, and analyzed by gel electrophoresis to quantify product formation of the lyase reaction. Another aliquot of preincubation mixture (200 μL) was injected into a 3 K, 100-500 μL dialysis cassette and dialyzed against 1X reaction buffer (2 x 1 L, 50 mM HEPES, 50 mM KCl, 5 mM $MgCl_2$, 200 mM EDTA, 0.01% Tween 20, pH 7.5) for 24 h at 4 °C. After dialysis, the lyase kinetics were analyzed following previously stated conditions.

Based on the evidence for irreversible inhibition and the nature of the DOB moiety to form crosslinks with Pol β, the interaction between **47** and Pol β was analyzed by mass spectrometry (MS) (Figure 37). A single modified peptide containing active site lysines K72 and K81 (Fragment 1) was detected upon trypsin digestion of intact Pol β incubated with **47** (10, 25 μM). MS/MS fragmentation was required to determine the modified residue(s). The relative intensity of the parent peptide was low and the only fragmentation observed contained unmodified K81 (y_3). By process of elimination, this suggested that K72 was modified.

Furthermore, in the 8 kDa domain incubated with **47** (10, 25 μM), two fragments containing K72 and K84 (Fragments 2 and 3) were detected with fragmentation data that supported K72 modification. The only internal lysine residue in Fragment 2 was K72. In addition, direct modification of K72 was observed in the MS/MS fragmentation (b_7*). A single modification was

also detected on Fragment 3. While there is no fragmentation data for peptide 3, we were confident modification occurred on K84 because it was the only internal nucleophilic residue.

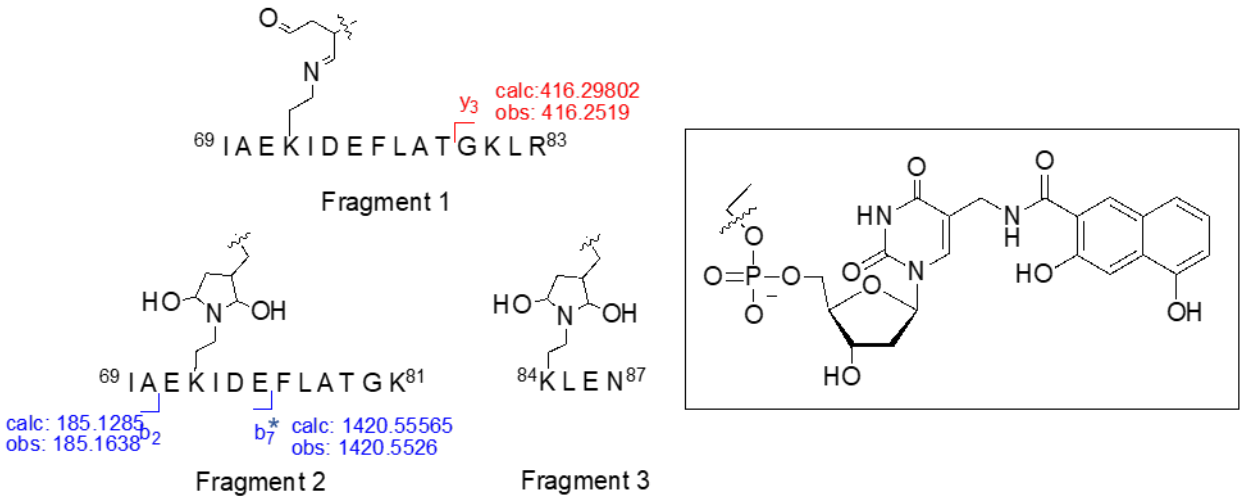


Figure 38. MS analysis of trypsin digestion of Pol β with **47.** Pol β (1 μM) and **47** (10 μM) were incubated in 1X reaction buffer (50 mM HEPES, pH 7.2, 5 mM $MgCl_2$, 2 mM DTT) at 25 ºC for 30 min. The mixture was concentrated with a 3K Amicon filter and washed 3-4 times with 1X reaction buffer. Digestion buffer (25 μL), 500 mM Tris HCl, pH 7.0), 10X trypsin (25 μL, 400 μM), and H_2O (175 μL) were added. The digestion sample (2 μM Pol β and 40 μM trypsin) was incubated at 37 ºC for 5 h. A portion (10 μL) of the digestion mixture was spun down and analyzed by UPLC-MS/MS using an ACQUITY UPLC HSS T3 Column (100 A, 1.8 μm, 2.1 mm x 100 mm). The flow rate was 0.3 mL/min running a gradient of 85:5:10 water: acetonitrile: 1% formic acid to 50:40:10 water: acetonitrile: 1% formic acid over 35 min. Analysis was conducted using BioPharmaLynx with tolerance set to 30 ppm and allowing for 4 missed cleavages.

We wanted to understand how inhibitor **47** selectively binds the 8 kDa domain but inactivated both the lyase and strand displacement activities in the holo-enzyme (Figures 30-36). By analyzing enzyme activity in the separate domains, we determined compound **47** had no effect on strand displacement synthesis by the 31 kDa domain (Figure 39), consistent with its binding in the 8 kDa domain.

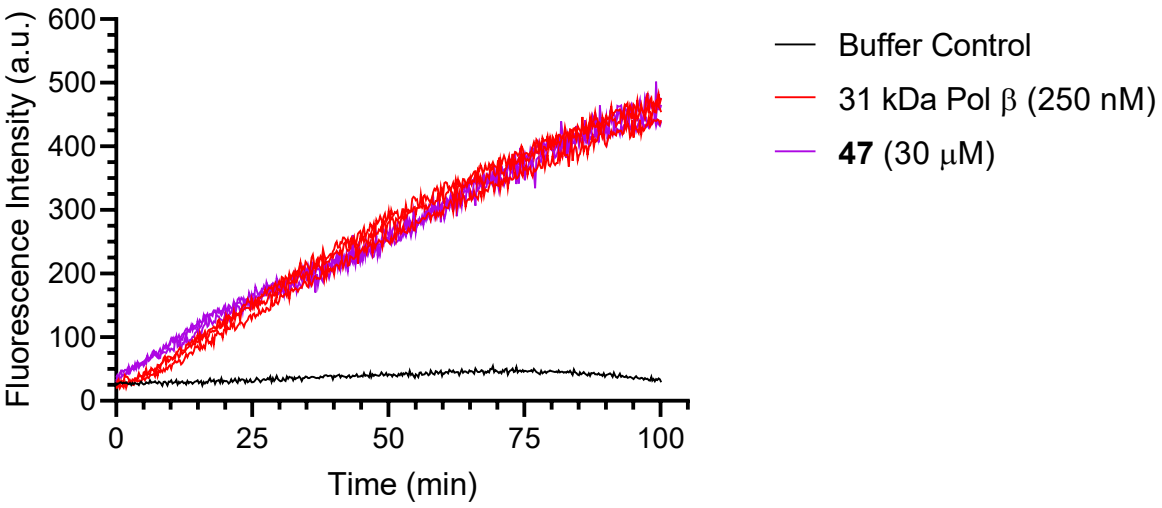


Figure 39. Inhibitor **47** did not inactivate strand displacement synthesis in the 31 kDa domain. 31 kDa Pol β (500 nM) was mixed with **47** (30 μM) for 30 min. Preincubation mixture was diluted 10-fold and mixed with 50 nM DNA substrate **B** and 100 μM dTTP.

To evaluate the relationship between domains and activities, we used fluorescence anisotropy to measure Pol β's DNA binding ability in response to inhibitor **47**. Fluorescence anisotropy monitors the tumbling speed of fluorescently labelled DNA (**F**, Table 1), which decreases in response to binding a large enzyme. We were unsure if the separate domains (i.e. 8 kDa, 31 kDa) would be large enough to distinguish a change in DNA tumbling rate. Despite the uncertainty, we found that compound **47** interfered with DNA binding in the holo-enzyme (Figure 40a) and the 8 kDa domain (Figure 40b) but had no effect on DNA binding in the 31 kDa domain (Figure 40c).

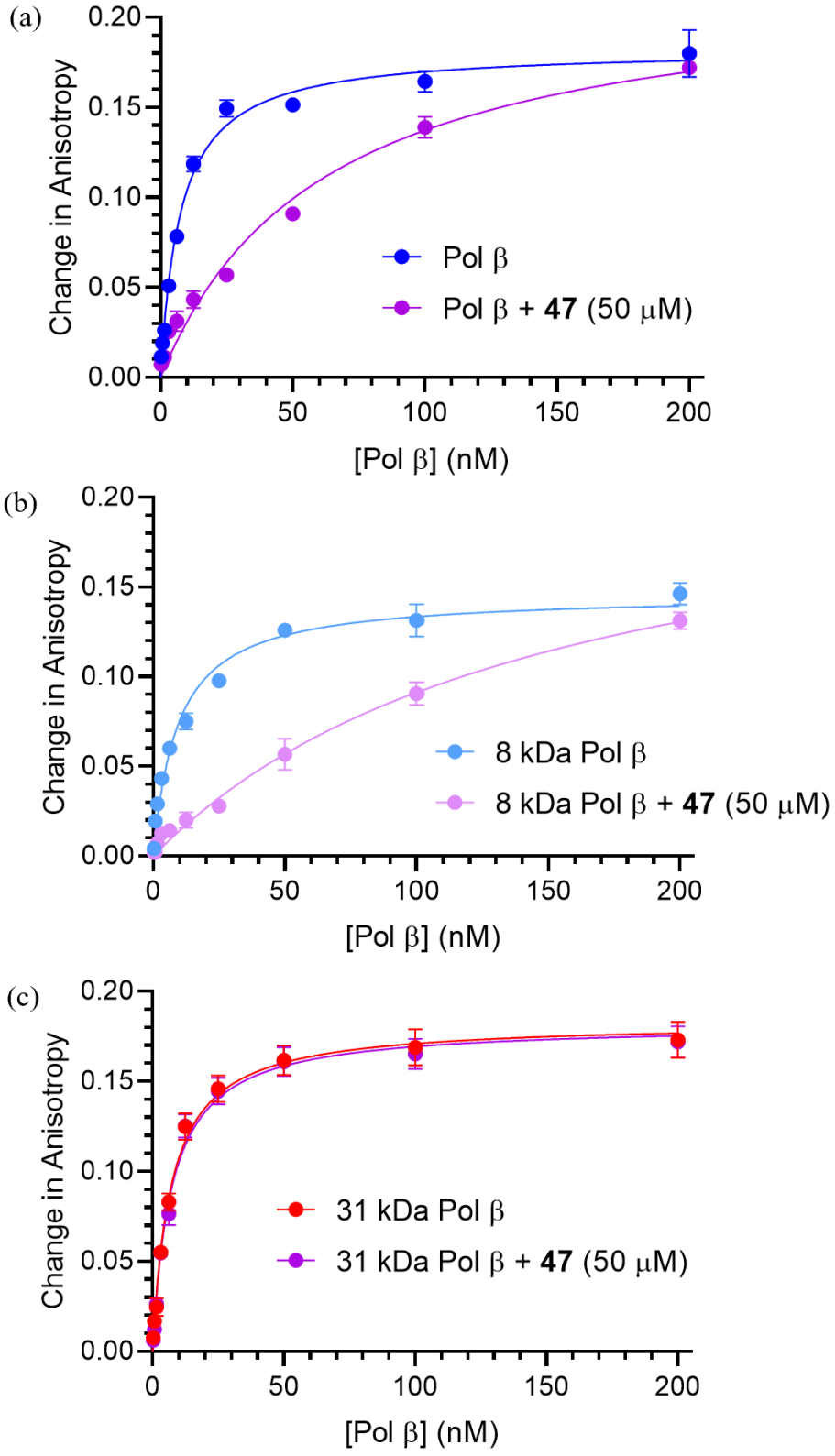
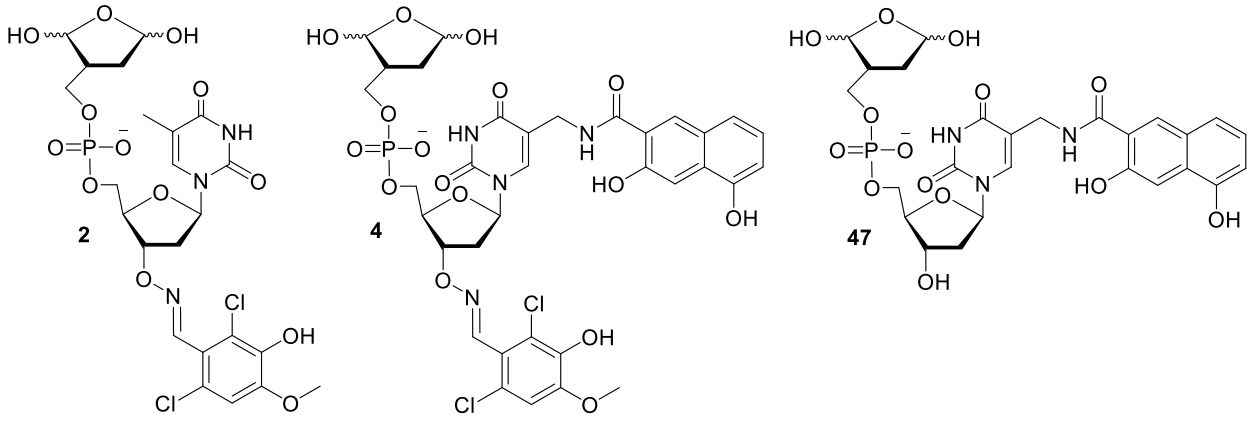


Figure 40. The effect of inhibitor **47** on DNA binding in WT Pol β (a), 8 kDa Pol β (b), and 31 kDa Pol β (c).

3.3.2 Polymerase domain selectivity

Intrigued by our previous results regarding 8 kDa domain specific binding, we were curious about previous Pol β inhibitors **2** and **4**.[26,27] Inhibitor **2** inhibited WT Pol β lyase activity with an IC_{50} of 21 μM and also inactivated polymerase activity.[26] At that time, the group assumed that these compounds targeted the lyase domain because DOB in DNA modified K72 (Section 2.3.1.2).[29] However, definitive proof to support this (i.e. the separate domains or MS) was unavailable. In addition, the structure of **4** resembled inhibitors **2** and **47**.



To accomplish this, we used our gel-based lyase assay with WT Pol β and the 8 kDa domain and our strand displacement activity with WT Pol β and the 31 kDa domain (Figure 41). We confirmed that inhibitor **2** inactivated the lyase activity only in the holo-enzyme and had little effect on the 8 kDa domain (Figure 41a). Conversely, inhibitor **2** interfered with strand displacement activity in both the intact enzyme and the 31 kDa domain (Figure 41b and c). These data indicated that inhibitor **2** binds in the polymerase domain but inhibits both activities.

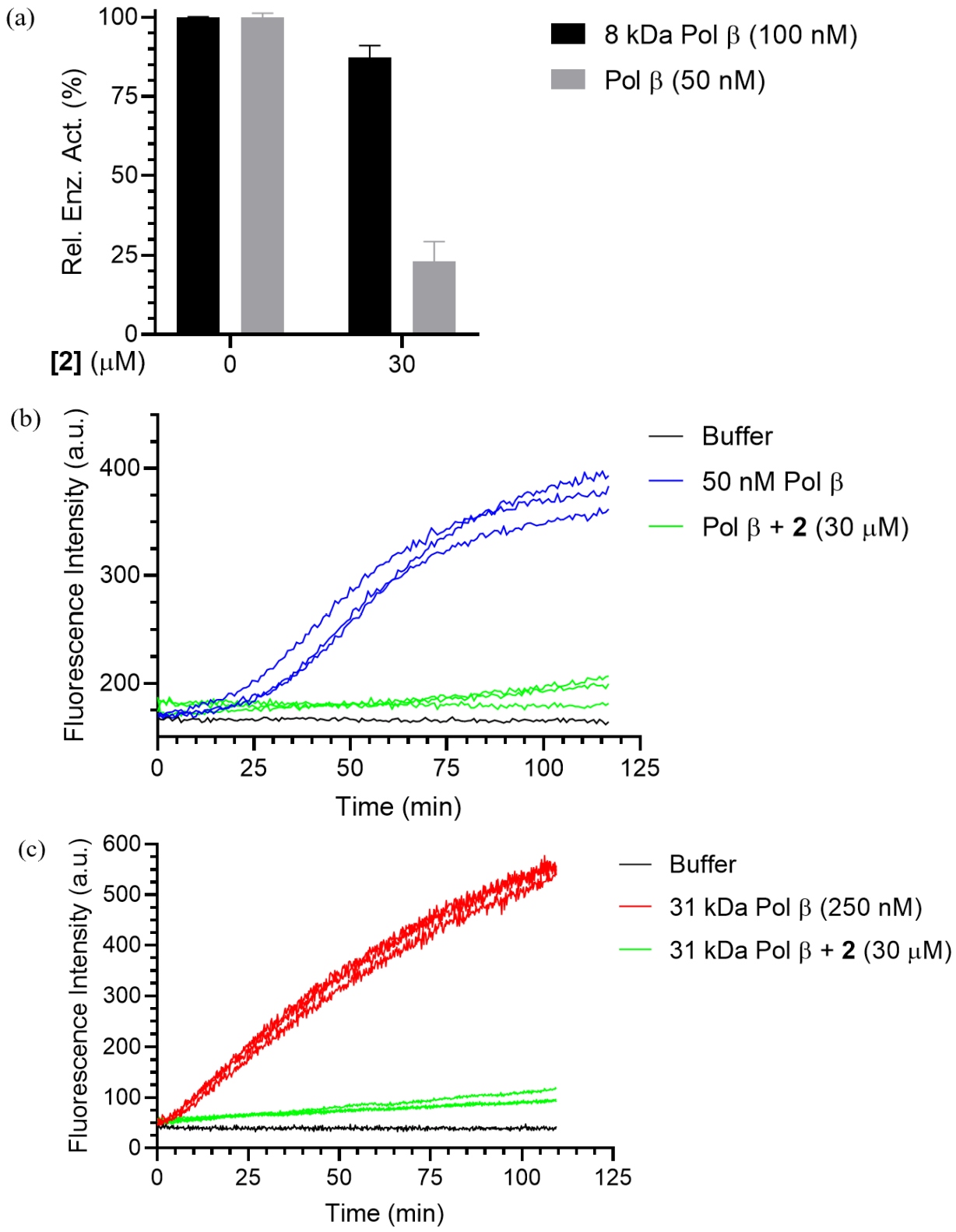


Figure 41. First-generation inhibitor (**2**) targeted the polymerase domain. (a) gel-based lyase assay with WT enzyme and 8 kDa domain, (b) strand displacement assay with WT enzyme, (c) strand displacement assay with 31 kDa domain.

While this result was not unwelcome, it was unexpected. As mentioned earlier, when in DNA, the DOB molecule modified Lys72 in the 8 kDa domain (Section 2.3.1.2).[29] Based upon this, the group was looking for lyase inhibitors and assumed that lyase inactivation also affected polymerase activity. The primary screening method detected compounds that inactivate strand displacement activity. Consequently, any lyase inhibitors that only inactivated lyase activity would not be identified in this screen.

We turned to fluorescence anisotropy to determine whether the original DOB-inspired inhibitor (**2**) also interfered with DNA binding in the holo-enzyme (Figure 42a) and the 31 kDa domain (Figure 42b). We analyzed fluorescence anisotropy in the intact enzyme with two concentrations of inhibitor **2** because we wanted to compare the effect on DNA binding with the same concentration used during kinetics (30 μM) and a higher concentration (60 μM) to ensure we could observe a significant difference in DNA binding. Sure enough, in the polymerase domain and the intact enzyme, DNA binding was disrupted in the presence of inhibitor **2**. The decrease in DNA binding in the holo-enzyme in response to inhibitor **2** was concentration dependent. Fluorescence anisotropy with the 8 kDa domain did not show any change in DNA binding in the presence of **2** (60 μM) (Figure 42c), consistent with the lack of lyase inhibition in the 8 kDa domain (Figure 41a).

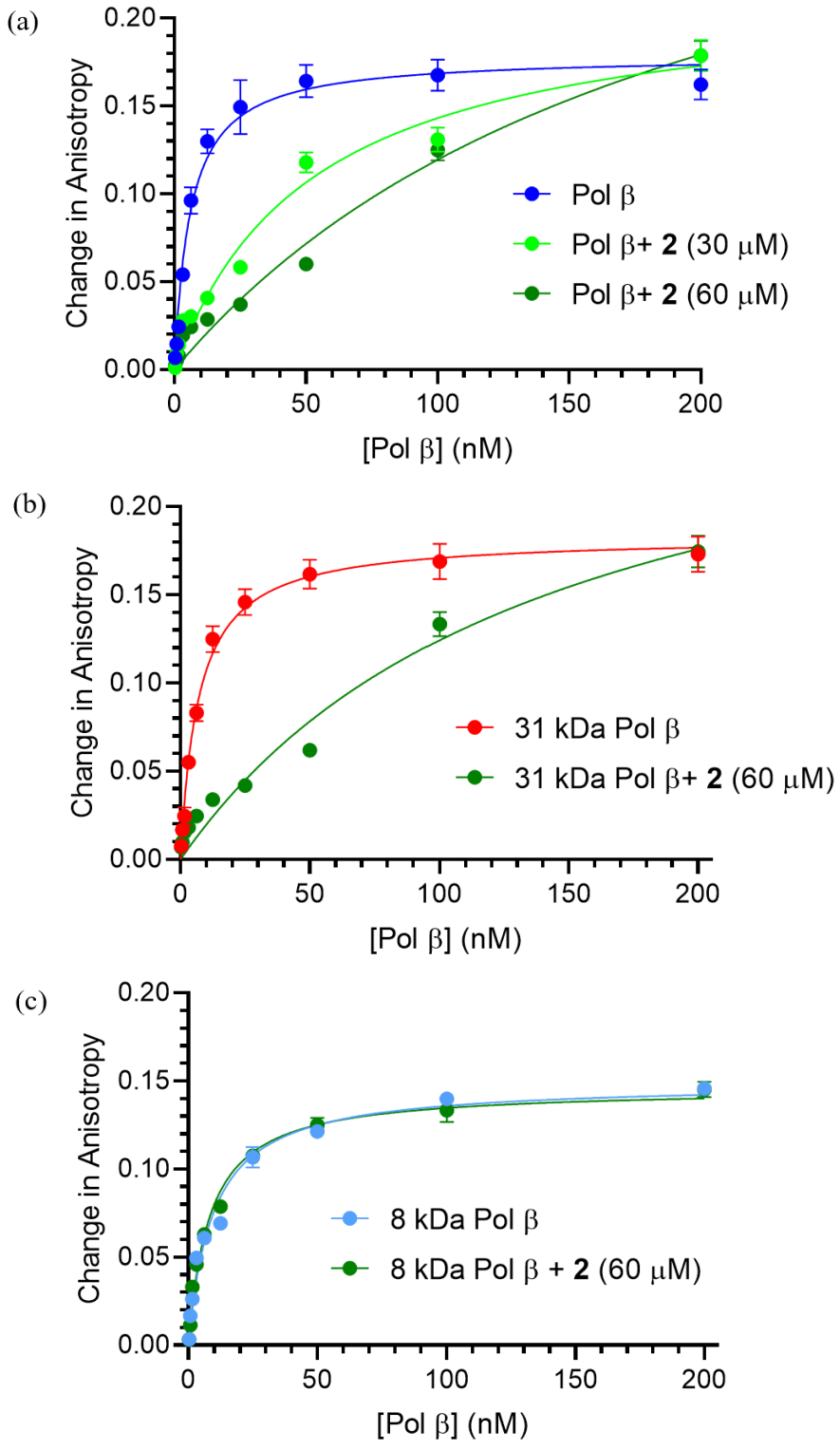


Figure 42. The effect of inhibitor **2** on DNA binding in WT Pol β (a) 31 kDa Pol β (b) and 8 kDa Pol β (c).

While separate domain kinetic assays are valuable for general binding information, they cannot determine which residue(s) was modified by the inhibitor. Therefore, we incubated Pol β with inhibitor **2** (5 μM), digested the protein with trypsin, and analyzed the peptide fragmentation by MS/MS (Figure 43). We detected several fragments but ultimately observed modifications on K127, K148, and K248. Fragments 4 and 5 overlap and cover amino acids 113-131. Fragment 6 contained two internal lysine residues (K113, K120, and K127). Fragmentation of this peptide (y_5* and y_7*) suggested that modification occurred between residues 125-131. By process of elimination, the only internal lysine between 125 and 131 was K127. Furthermore, the only internal lysine in Fragment 5 was K127. Together, these indicated covalent modification occurred on K127.

Fragments 6 and 7 span amino acids 138-152. In this region, K142 and K148 were present but fragmentation in Fragment 6 suggested K148 is the site of modification. In Fragment 6, K142 was detected unmodified (b_6) and K148 was detected modified (y_6* and y_9*). Additional evidence for modification on K148 appeared in the detection of Fragment 7, where K148 was the only internal lysine available for covalent modification.

Fragments 8-10 overlap and consist of amino acids 235-258. Internal lysines included K244 and K248. In Fragment 8, K244 was unmodified (b_{10}) but the longer fragmentation consisting of both K244 and K248 (b_{17}*) appeared modified. Fragments 9 and 10 directly pointed to modification on K248 as it was the only internal lysine. Furthermore, these fragments start at amino acid 245, which means trypsin would have cleaved the protein at K244, suggesting it was unmodified. Fragmentation data of these peptides also supported modification on K248 (b_6*, y_{11}*, Fragment 9; b_6*, y_7*, y_8*, y_9*, Fragment 10). It is interesting that both the hydrated and dehydrated versions of the modification were observed across detected fragments.

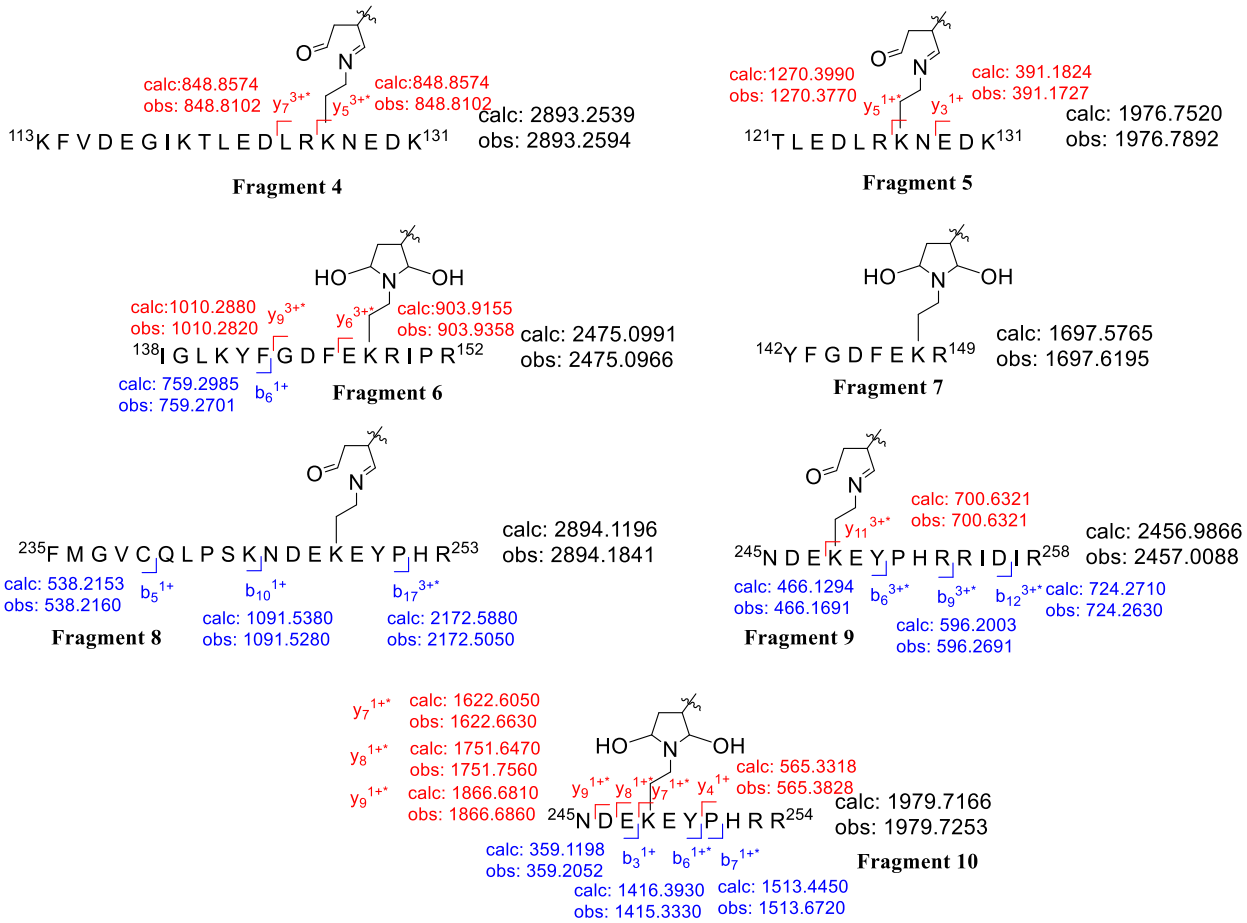


Figure 43. MS analysis of trypsin digestion of Pol β incubated with **2**. Pol β (1 mM) and **2** (5 μM) were incubated in 1X reaction buffer (50 mM HEPES, pH 7.2, 5 mM $MgCl_2$, 2 mM DTT) at 25 ºC for 30 min. The mixture was concentrated with a 3 K Amicon filter and washed 3-4 times with 1X reaction buffer. Digestion buffer (25 μL), 500 mM Tris HCl, pH 7.0), 10X trypsin (25 μL, 400 μM), and H_2O (175 μL) was added. The digestion sample (2 μM Pol β and 40 μM trypsin) was incubated at 37 ºC for 24 h. A portion (10 μL) of the digestion mixture was spun down and analyzed by UPLC-MS/MS using an ACQUITY UPLC HSS T3 Column (100 A, 1.8 μm, 2.1 mm x 100 mm). The flow rate was 0.3 mL/min running a gradient of 85:5:10 water: acetonitrile: 1% formic acid to 50:40:10 water: acetonitrile: 1% formic acid over 35 min. Analysis was conducted using BioPharmaLynx with tolerance set to 30 ppm and allowing for 4 missed cleavages.

After MS analysis, we located these residues in the crystal structure of Pol β (Figure 44). We were surprised to find that none of the detected modified residues appear near the DNA, and yet we know that inhibitor **2** interferes with DNA binding (Figure 42). Furthermore, the structure of the polymerase domain does not change significantly upon binding the DNA.[114,116] The position of lysines slightly shift position upon DNA binding (Figure 44).

It is possible that modification at these lysines impact folding or initiate a conformational change that disrupts DNA binding and activity. However, it is likely that the discrepancy observed is just a matter of concentration of the inhibitor. The concentration of **2** (5 μM) needed for MS analysis is significantly lower than the concentration used (30-60 μM) in fluorescence anisotropy experiments. It was possible that inhibitor **2** was not very selective and that at higher concentrations, it would modify several lysines that, cumulatively, inhibit DNA binding. In support of this, we digested Pol β incubated with **2** (25 μM) and observed modifications on several other lysines (i.e. K84, K113, K127, K167, K148, K234, K248), one of which has been implicated in DNA binding interactions (i.e. K234).[117]

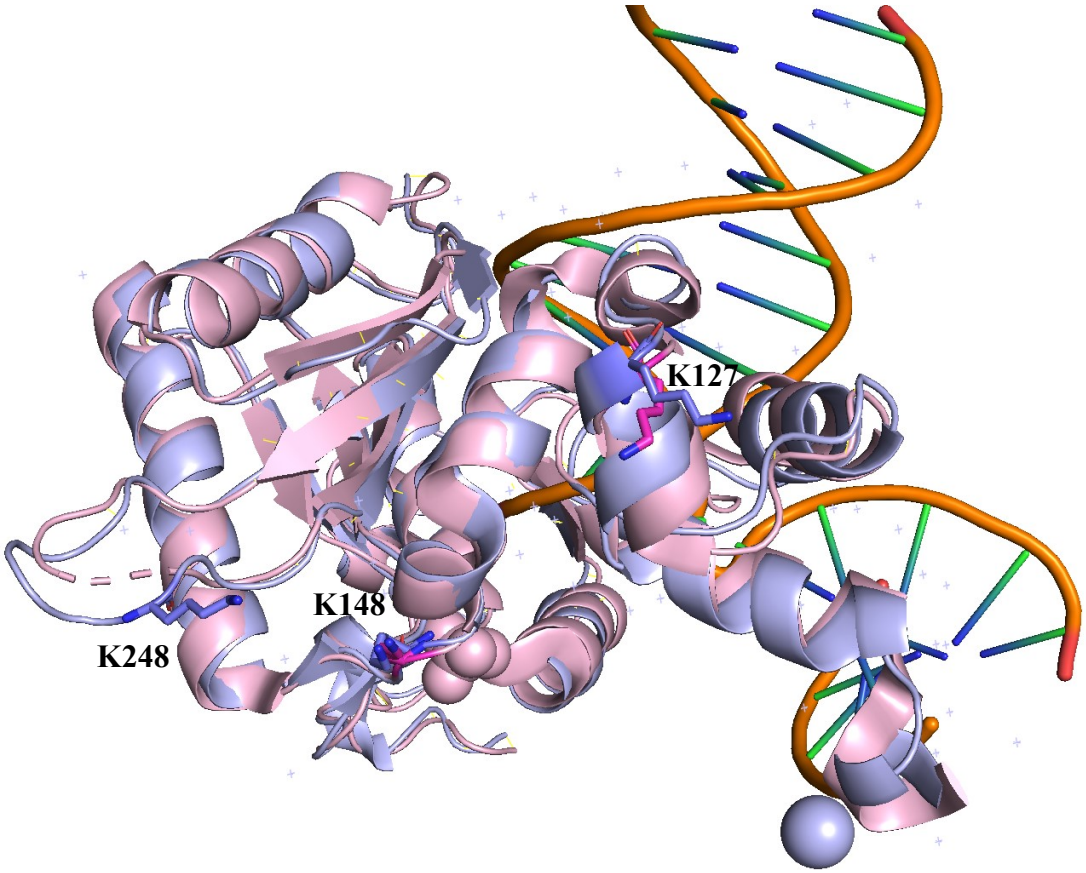


Figure 44. Conformational changes when Pol β binds to DNA. Key: pink = 31 kDa Pol β without DNA (PDB: 1BPD), purple = 31 kDa Pol β with DNA (PDB: 1BPX), sticks = modified lysine residues.

It is worth noting that just because inhibitor **2** binds in the polymerase domain, it is unclear whether second-generation inhibitor **4** does. This is unclear because **4** contains components of inhibitor **2** and **47**, but these individual inhibitors bind the 31 kDa polymerase and 8 kDa lyase domains respectively (Table 2).

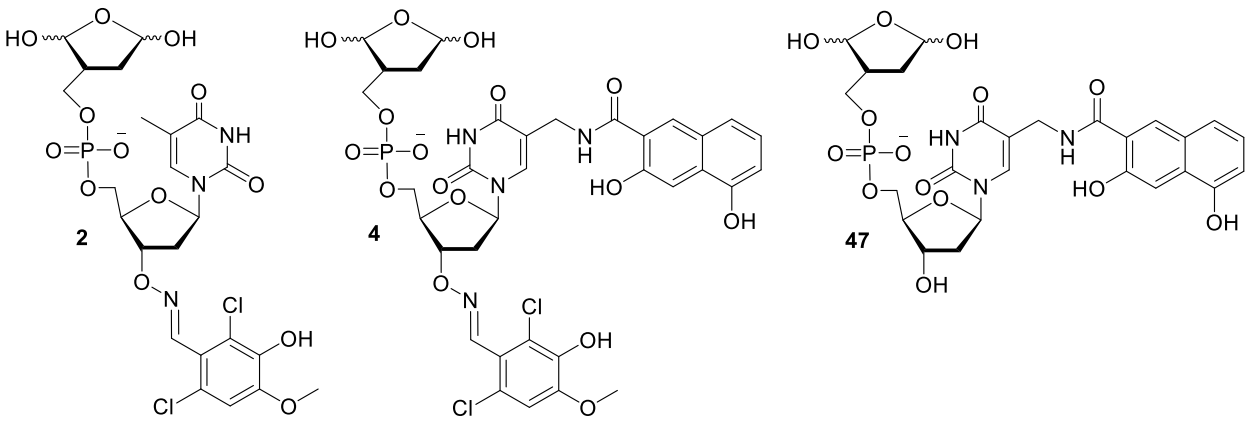


Table 2. Summary of inhibitor binding locations and activity inhibition.

	2	**4**	**47**
Targets	31 kDa	ND	8 kDa
Lyase Activity	Inhibits	Inhibits	Inhibits
Polymerase Activity	Inhibits	Inhibits	Inhibits
MS	K127, K148, K248	ND	K72, K84
Disrupts DNA binding	Y – WT, 31 kDa N – 8 kDa	ND	Y – WT, 8 kDa N – 31 kDa

Y = Yes, N = No

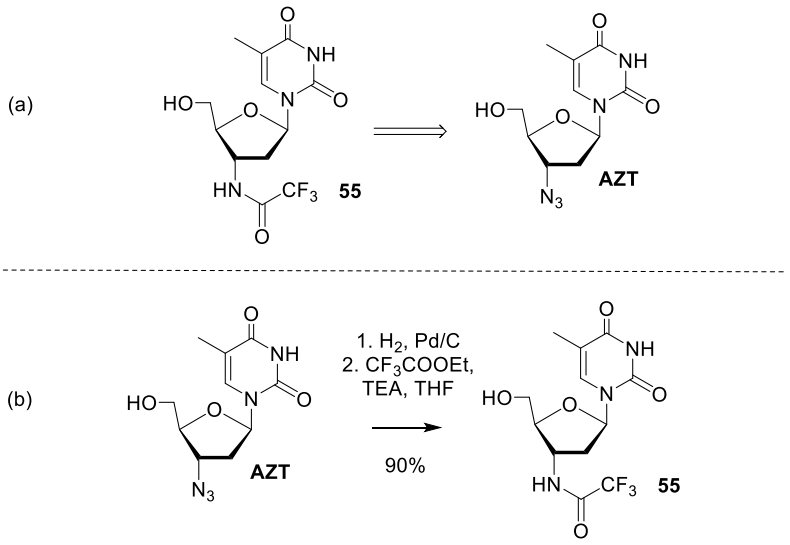
The relationship between location of inhibitor binding and the activity(ies) inhibited is still unclear. How do inhibitors **2** and **47** bind different domains but inactivate both activities? We hypothesize that, in addition to the disruption of DNA binding, in certain cases, an allosteric conformational change occurs that interferes with the other domain DNA binding and activity. However, the current experiments are inadequate for correlating where and how an inhibitor binds with an allosteric effect. We imagine that solution NMR would be capable of identifying protein

dynamics and conformational changes that occur in response to inhibitor binding. In addition, X-ray crystallography could be used to observe an allosteric Pol β conformational change. Regarding this second approach, inhibitor **47** has been sent to a collaborator for X-ray analysis.

3.3.3 The Significance of the 3'-oxime and the 3'-amide

After identifying inhibitor **47**, we altered the synthesis of the 3'-functionalized library. This synthesis started with commercially available 3'-azidothymidine (AZT) and after reduction and *N*-trifluoroacetamidation yielded **55** (Scheme 26). Starting with AZT significantly simplified the synthesis by avoiding consecutive Mitsunobu reactions. The Mitsunobu reaction was necessary when starting from thymidine to ensure stereochemistry retention at the C3'-position (Scheme 10).

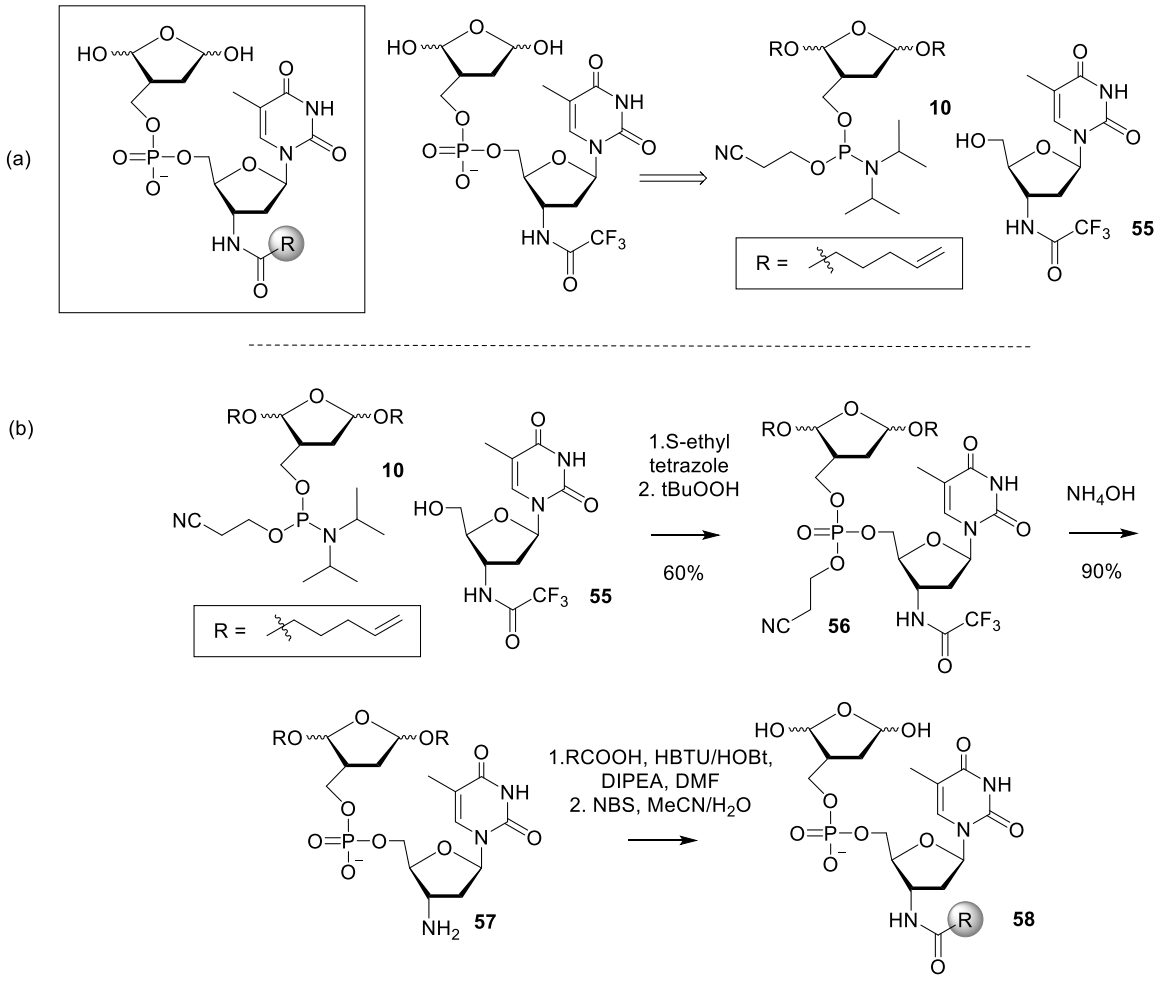
Scheme 26. Retrosynthetic (a) and forward (b) synthesis of 3'-*N*-trifluoroacetamide containing monomer **55**.



The first-generation scaffold was prepared by coupling compound **55** and **10** (Scheme 12) to generate the phosphate triester (**56**) (Scheme 27). Scaffold (**57**) was isolated following deprotection with concentrated aqueous ammonia and was used to generate library **58** (Section

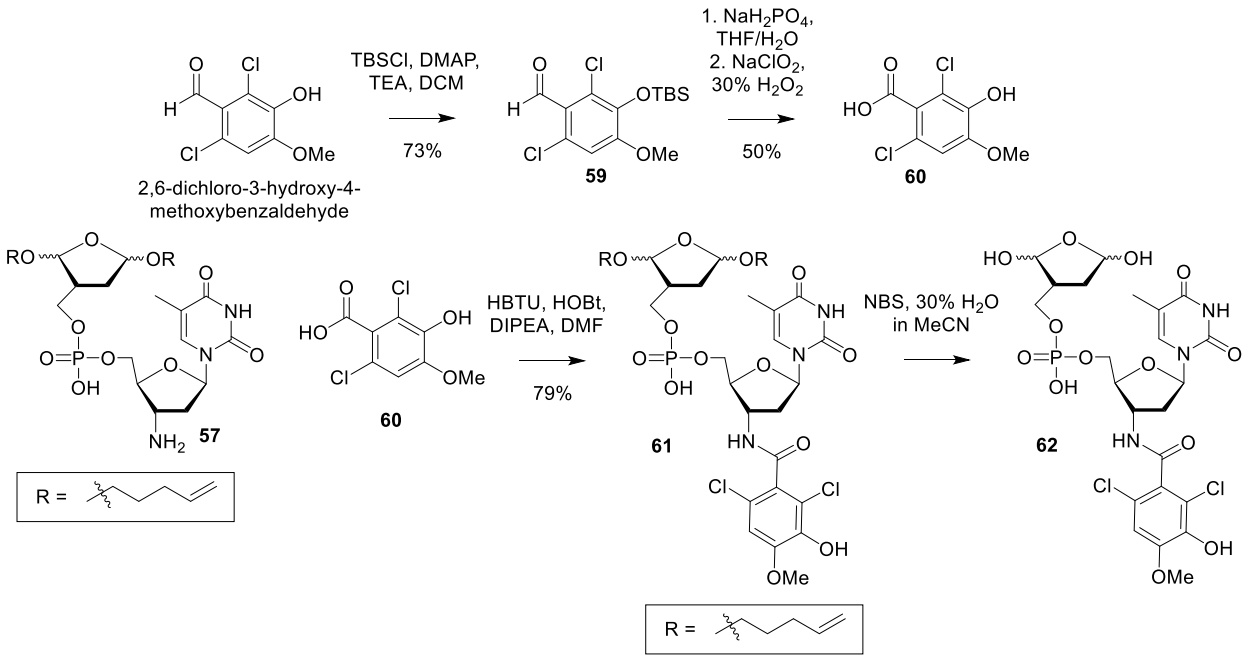
3.4). However, we wanted to explore the impact of changing this functional group in the context of our inhibitors. In library **1**, the 3'-alkoxyamine was functionalized by a library of aldehydes to form oximes (Figure 21). In contrast, library **58** consisted of a 3'-amine that was functionalized by a library of carboxylic acids to yield amides. The oxime and the amide are different in flexibility and length. These differences position the aryl group in different orientations and the resulting molecules can, therefore, experience unique binding interactions.

Scheme 27. Retrosynthetic (a) and forward (b) synthesis of 3'-amine scaffold (**57**).

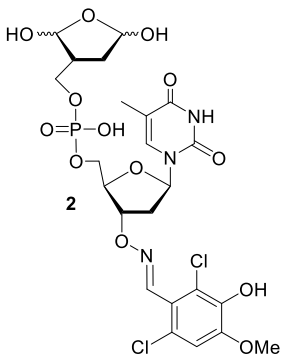


To evaluate whether this difference affect enzyme inhibition, we resynthesized inhibitor **2** but in the context of the amide bond. In other words, we used the same commercially available aldehyde used in the synthesis of inhibitor **2** and transformed it into the corresponding carboxylic acid (Scheme 28). The 3'-alcohol was protected with a TBS group to yield **59**. The carboxylic acid (**60**) was obtained in the subsequent oxidation, in which the TBS group fell off during the reaction and work up.[164] Our optimized coupling conditions (i.e. HBTU and HOBt, Section 3.3) were used to couple scaffold **57** and **60** to yield **61**. The pentenoyl groups were cleaved under oxidative conditions to yield the final compound (**62**) bearing the same aryl group in inhibitor **2**, confirmed by MS.

Scheme 28. Synthesis of amide version of **2** (**62**).



Compound **62** was analyzed using the strand displacement assay at multiple concentrations (Figure 45). It was clear that at concentrations as high as 50 μM, **62** did not inhibit WT Pol β. This contrasted with inhibitor **2**, which has an IC_{50} of 21 μM. Based on this experiment, we concluded that the differences between the oxime and the amide group were significant enough to yield different inhibitor hits. It is likely that the additional flexibility and length of the oxime positioned the substituted benzyl group uniquely. This result was reassuring because it indicated that the libraries are producing molecules that make specific interactions with the enzyme.



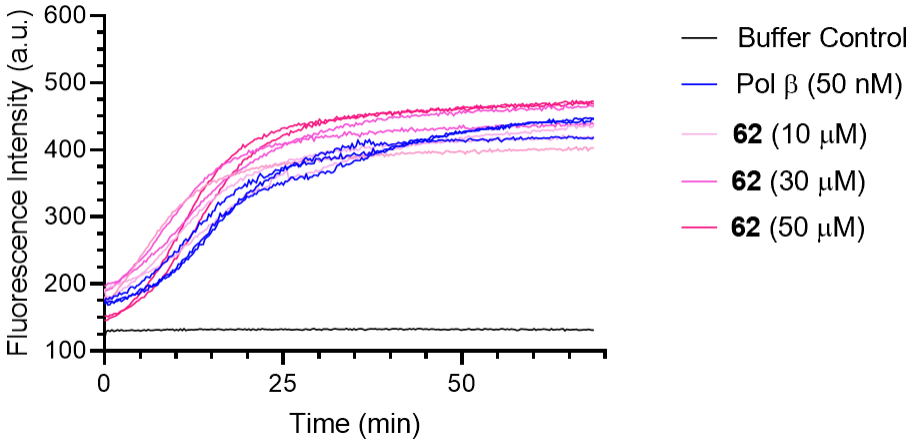


Figure 45. Compound **62** does not inhibit strand displacement activity of the intact enzyme. Pol β (50 nM) was preincubated with **62** at various concentrations (listed) for 30 min. The preincubation mixture was diluted 10-fold and mixed with DNA cocktail solution, to yield 5 nM Pol β, 50 nM DNA substrate **B,** and 100 μM dTTP**.**

3.4 WT Pol β inhibitor(s) starting from AZT

While inhibitor **47** provided a useful lead, the development of the 3'-library route through AZT (Schemes 26 and 27) was more convenient and efficient. Therefore, scaffold **57** was used to create a first-generation library (**58**). A modest polymerase inhibitor was identified in library **58**

(Section 3.4.1). This hit compound led to the design of a second-generation library, where a potent and selective Pol β inhibitor was identified (Section 3.4.2). This compound was characterized extensively via *in vitro* (Section 3.4.2.1) and *in vivo* (Section 3.4.2.3) assays.

3.4.1 First-generation WT Pol β inhibitor: Functionalization at the 3'-position

In addition to modification at the 5-methyl position of thymine (Section 3.3.1), a first-generation library involving 3'-functionalization was synthesized (Scheme 27, Section 3.3.3). Scaffold (**57**) (40 mg total, 200 nmol per carboxylic acid) was coupled to 280 carboxylic acids (Appendix 1) and deprotected with NBS to form the 3'-functionalized library (**58**). This first-generation library was screened using the same conditions previously mentioned (Section 3.3.1).

Library **58** led to the identification of hit compounds **63** and **64** (Figure 46). First round hits were screened multiple times to rule out false positives. Inhibition by hit compounds were compared to the inhibitory effects of the free acids. After three rounds, no other inhibitors were consistently identified in the library. Strand displacement inhibition was dependent on the concentration of **63** and **64** (red and green, Figure 47). By comparing the increase in fluorescence in response to inhibitor concentration, compound **63** (30 μM) completely inactivated the strand displacement, while the same concentration of **64** (30 μM) resulted in much less activation (Figure 47).

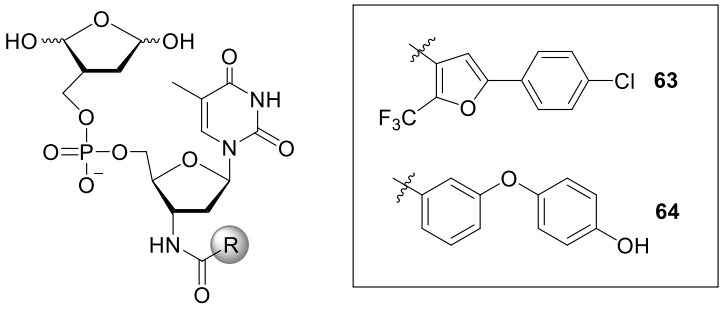


Figure 46. Strand displacement screen identified Hit compounds **63** and **64** against Pol β.

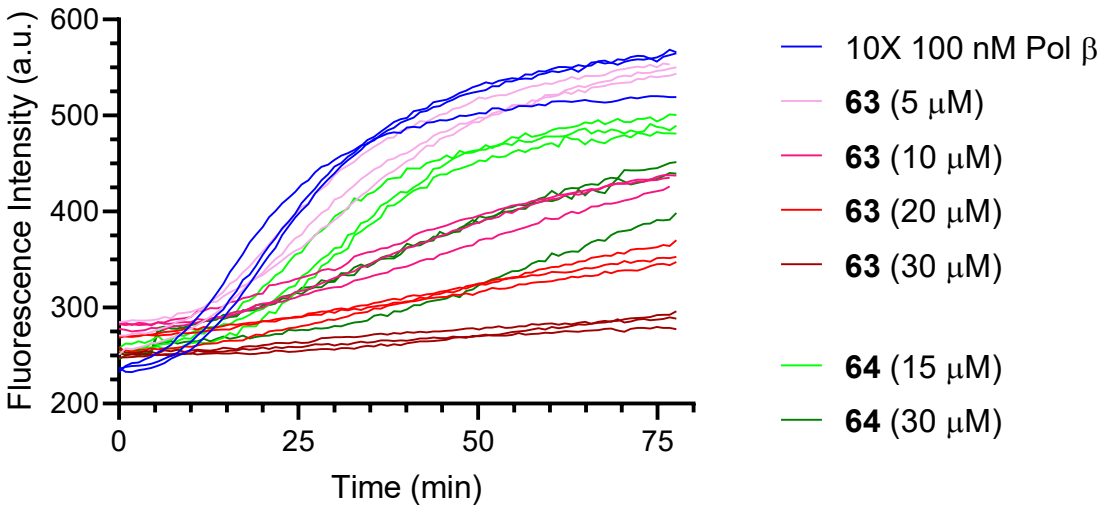


Figure 47. Pol β strand displacement inhibition is dependent on the concentration of **63** and **64**. Pol β (100 nM) was preincubated with hit **63** or **64** at various concentrations (listed) for 20 min. The preincubation mixture was diluted 10-fold and mixed with DNA cocktail solution, to yield 5 nM Pol β, 50 nM DNA substrate **B,** and 100 μM dTTP**.**

The quantitative gel-based lyase assay indicated that hits **63** and **64** did not inactivate the lyase activity even at much higher concentrations (Table 3). This suggested that **63** and **64** did not inhibit the lyase activity. This result was unexpected. As mentioned earlier, while we identified Pol β inhibitors that selectively bind distinct domains (i.e. **2** and **47**), they inactivated both activities (Section 3.3). Regardless, **63** was clearly a more potent polymerase inhibitor compared to **64** (Figure 47). Therefore, **63** was analyzed further using the separate domains of Pol β. Hit compound **63** had no impact on the 8 kDa lyase domain activity but inhibited the strand displacement activity of the 31 kDa polymerase domain at 15-20 μM (Figure 48). The observations were consistent with those described above with intact Pol β, which suggested that **63** targets the polymerase domain. Hit **63** was distinct from previous inhibitors (**2**, **47**) because it selectively targeted the 31 kDa domain and inactivated the polymerase activity, without affecting lyase activity.

Table 3. Summary of Pol β lyase inactivation by **63** and **64**.

Hit	Concentration (μM)	Relative lyase activity (%)
63	25	75 ± 11
	30	87 ± 9
	60	46 ± 5
	120	46 ± 7
64	25	95 ± 5

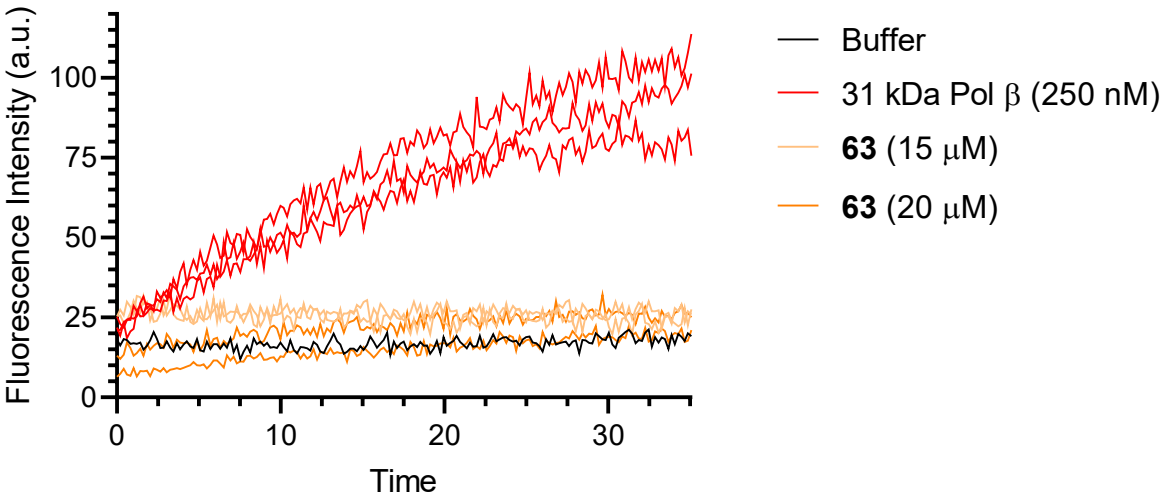


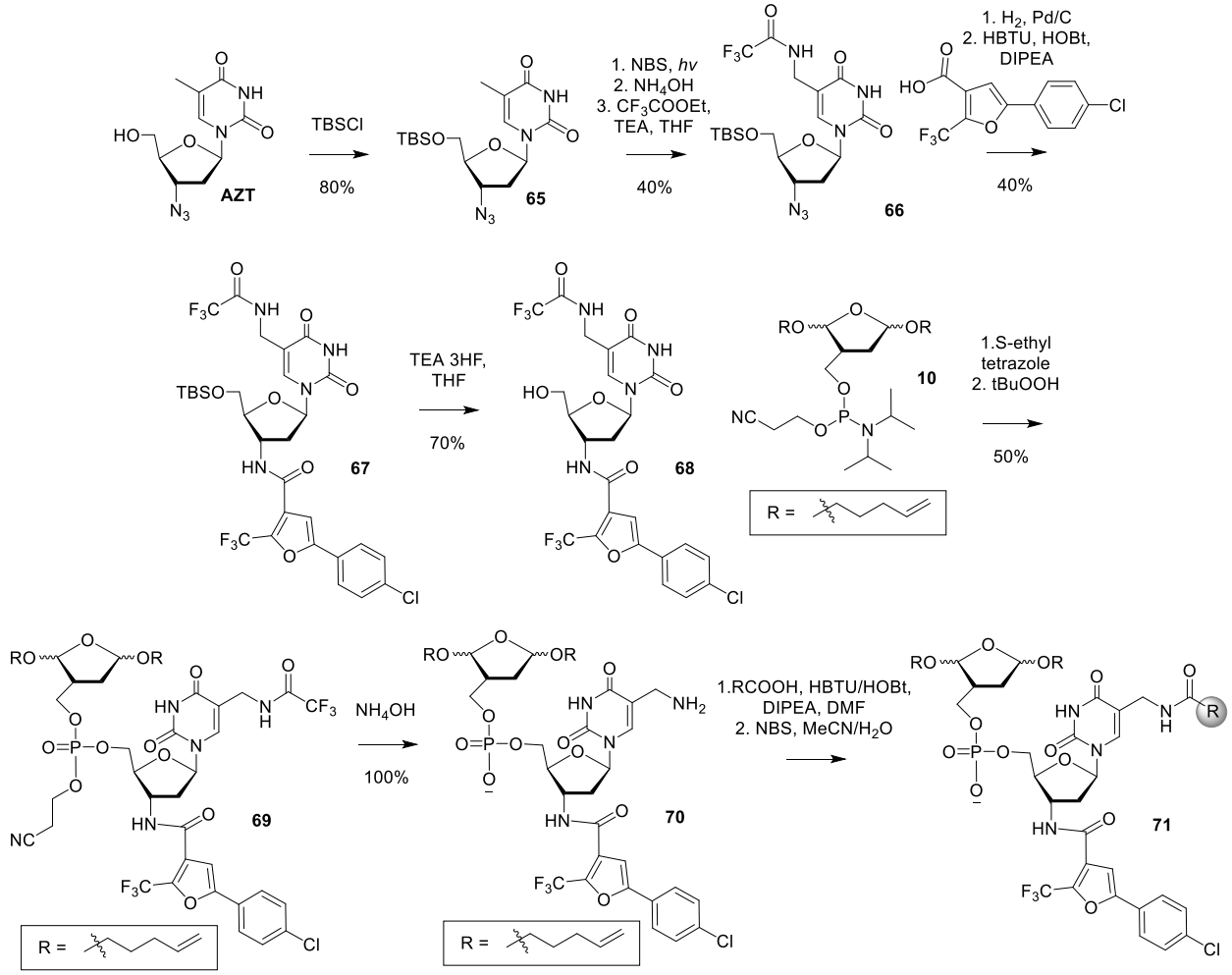
Figure 48. Inactivation of the Pol β 31 kDa domain strand displacement activity. 31 kDa Pol β (500 nM) was mixed with **63** (15 or 20 μM) for 30 min. Preincubation mixture was diluted 10-fold and mixed with 50 nM DNA substrate **B** and 100 μM dTTP.

3.4.2 Second generation inhibitor: Functionalization of 59 at the thymine 5-methyl position

The synthesis of the 3'-first-generation library (**58**) was efficient, effective, and identified an interesting hit candidate (**63**). Therefore, hit compound **63** guided the synthesis of a second-generation library (Scheme 29). The second-generation library synthesis started with the silylation of AZT (**65**). Following a one pot reaction including bromination, ammonolysis and *N*-trifluoroacetamidation (**66**), the 3'-azide was reduced and coupled to the carboxylic acid in the previous first generation hit compound (**63**) to yield intermediate **67**. This intermediate underwent

desilylation (**68**) and coupling with **10** (Scheme 12) to yield phosphate triester **69**. The scaffold (**70**) was generated after the β-cyano ethyl and *N*-trifluoroacetamide groups were removed under basic conditions. A library of carboxylic acids (~325 compounds) was coupled to the 5-aminomethyl position (**70**, 25 mg total, 100 nmol per acid) and the reactive DOB moiety was deprotected with NBS to generate the second-generation library (**71**).

Scheme 29. Synthesis of a second-generation library containing the 3'-hit compound moiety.



In the second-generation screen, each compound (700 nM) was preincubated with Pol β (50 nM) for 25 min, mixed with a 5X cocktail solution containing dTTP and fluorescently labeled ternary complex (Figure 24, **B**, Table 1), and diluted together in 1X buffer (10X for

preincubation, 5X for DNA). The concentration of library compound(s) during the screen was much lower than the concentration in the first-generation screen. Higher concentrations (5 μM) led to the identification of a large portion of the library (~50% of the compounds tested) as potential hits. This was encouraging because it suggested that we were going to identify a potent inhibitor. Even screening at a much lower concentration, the initial screen identified multiple compounds, so all hits were resynthesized and rescreened along with the free acids to rule out false positives (Section 3.3.1). After three rounds of screening, this library led to the identification of inhibitor hit (**72**, Figure 49). No other compounds with similar potency were identified using these screening conditions. This inhibitor hit bore a 3-hydroxylated naphthoic acid moiety, a structural element present in several other inhibitors libraries (i.e. **47** and **48**) and inhibitors (i.e. **4** and PA).[27,145]

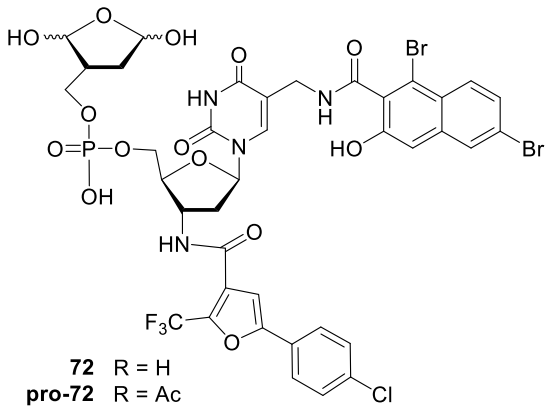


Figure 49. Hit compound and pro-inhibitor (**72**).

Compound **72** did not significantly inhibit the lyase activity of Pol β (Figure 50). At 5 μM **72**, Pol β retained 75% of its lyase activity. This was not surprising considering that it was derived from the first generation hit **63**, which did not inhibit the lyase activity. However, this meant that the gel-based lyase assay would not be applicable when trying to determine IC_{50} values and other quantitative data for inhibitor **72**.

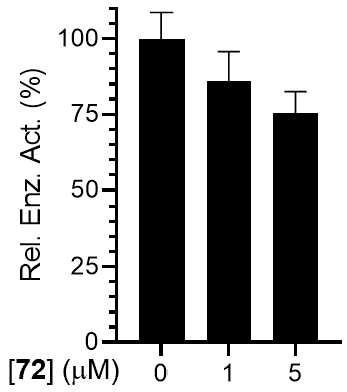


Figure 50. Compound **72** did not inhibit lyase activity.

Consequently, the strand displacement assay (Figure 24) was used to extract kinetic information (Section 3.2.1). The irreversible inactivation of Pol β depends on the concentration of **72** and preincubation time (Figure 51). Kinetic parameters of Pol β inactivation by **72** were gleaned from the data, including an IC_{50} and a K_I and k_{inact} (Figures 52 and 53, and Table 4).[159]

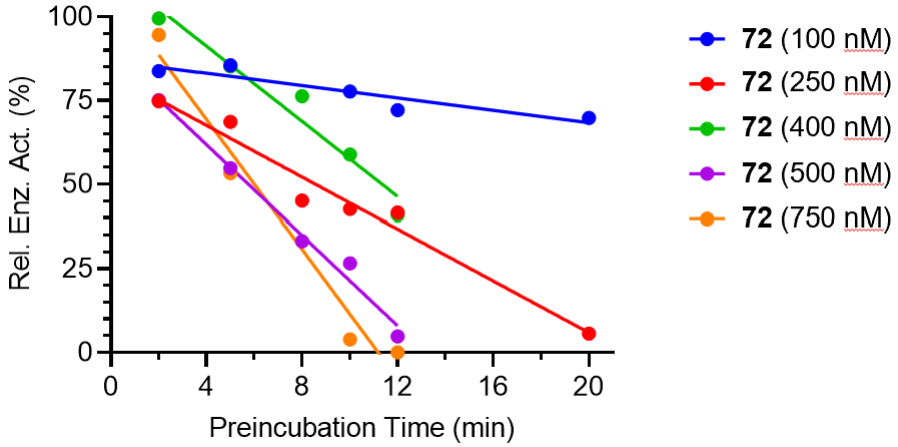


Figure 51. Pol β strand displacement inhibition is dependent on concentration and preincubation time with **72**. Pol β (50 nM) was preincubated with **72** at various concentrations for multiple preincubated times. After preincubation time, the mixture was diluted 10-fold and mixed with 30 nM DNA substrate **B** and 300 μM dTTP. Relative rate of 50 nM Pol β in the presence of **72** at various concentrations (listed) compared to the rate of 50 nM Pol β in the absence of inhibitor.

An IC_{50} value is not an appropriate method for evaluating an irreversible inhibitor because they are dependent on preincubation time (Section 3.2.2). An IC_{50} plot looks at a single

preincubation time and can, therefore, be misleading. This was evident in Figure 52 which displays two curves from different preincubation times that yield different IC_{50} values. With a 10 min preincubation time, the IC_{50} value of **72** is 458 nM. However, a 12 min preincubation time yields an IC_{50} value of 409 nM.

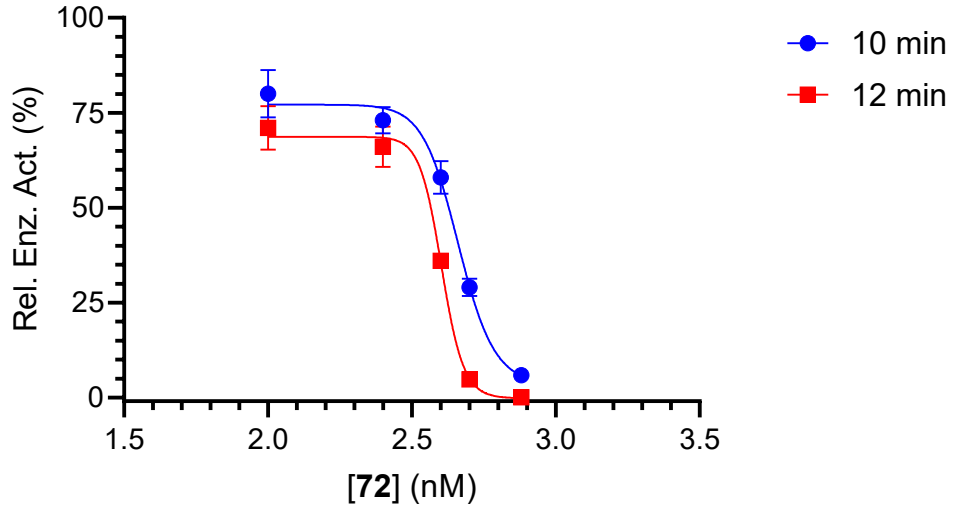
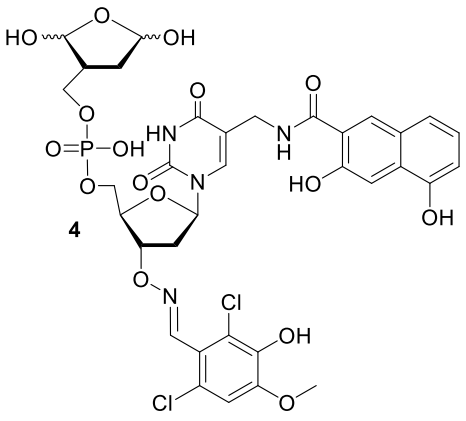


Figure 52. IC_{50} plots for **72** at two preincubation times. Pol β (50 nM) was preincubated with **72** at various concentrations for 10 min (blue) or 12 min (red). After preincubation time, the mixture was diluted 10-fold and mixed with 30 nM DNA substrate and 300 μM dTTP.

Although an IC_{50} value is not preferred to other kinetic measurements, it is the only method available to compare **72** to previous second-generation inhibitor **4**.[27] However, the conditions used thus far to evaluate the IC_{50} of **72** were different from those used for **4**. The differences include shorter preincubation time and greater dilution for **72**. Both factors are more stringent conditions for the inhibitor and therefore, led us to think that **72** may be more potent than **4** (Section 3.3). Therefore, the dilution and preincubation conditions used to evaluate **4** were repeated for **72** (Table 4, Appendix Figure 1). The IC_{50} value for **72** was 204 nM



against Pol β, compared to **4**, which had an IC_{50} value of 400 nM against Pol β.[27] It should be noted that the IC_{50} value of 204 nM for **72** cannot be compared directly to the IC_{50} values obtained from 10 and 12 min preincubations (separated by a thick bar in Table 4) because the preincubation dilutions were also different (10 and 12 min, 10X; 30 min, 1.1X).

Table 4. Summary of kinetic parameters of **72** inhibition of Pol β.

	k_{inact} (s^{-1})	K_I (μM)	k_{inact}/K_I ($M^{-1}s^{-1}$)	IC_{50} (nM)	Preincubation time (min) (for IC_{50} only)
72	$7.0 \pm 1.0 \times 10^{-3}$	1.8 ± 0.45	3.89×10^{3}	458	10
				409	12
				204	30
4	-	-	-	400	30

Based on the previous discussion, an IC_{50} value differed depending on whether it was derived from a shorter or longer preincubation time. As a result, inhibitor kinetic parameters, such as K_I and $k_{inact,}$ are more meaningful. A Kitz- Wilson plot was used to determine these kinetic constants (Section 3.2.2).[159] The K_I of **72** is 1.8 ± 0.45 μM and the k_{inact} is $7.0 \pm 1.0 \times 10^{-3}$ s^{-1} (Figure 53, Table 4).

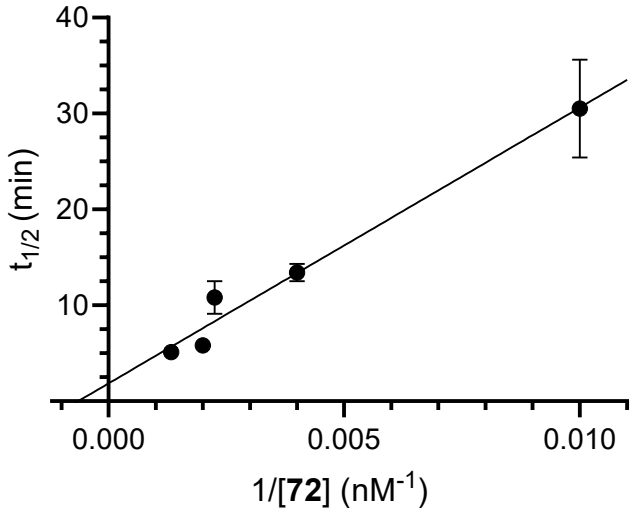


Figure 53. Kitz-Wilson plot of **72**.

The K_I represents inhibitor binding affinity and provides insight about the dosage needed to inhibit 50% of the enzyme target. Therefore, a smaller K_I value indicates greater binding and a smaller concentration required to inhibit the target. A large k_{inact} value is more desirable because it indicates a quick rate of enzyme inactivation. Furthermore, the k_{inact}/K_I is a measure of overall inhibitor efficiency. Covalent inhibition is becoming increasingly more popular. Unfortunately, most potent, irreversible inhibitors of Pol β do not report these kinetic parameters so a direct comparison cannot be made.[26,27] Recently, a low potency Pol β lyase inhibitor (K_I = 1.44 mM) was identified whilst searching for compounds that directly engage with the lyase active site lysine (i.e. K72).[165] Alternatively, while some irreversible inhibitors exist that target other polymerases, their kinetic parameters are limited to IC_{50} values.[166,167] Instead, we must compare **72** to irreversible inhibitors of other targets. For comparison, a novel covalent inhibitor of the histone methyltransferase nuclear receptor-binding SET domain protein 1 (NSD1) exhibited K_I = 25.0 ± 5.2 μM, k_{inact} = 4.8 ± 0.5 x 10^{-4} s^{-1}, and k_{inact}/K_I = 19.6 $M^{-1}s^{-1}$.[168] Despite the difference in target, inhibitor **72** exhibited a greater binding affinity for Pol β, a faster rate of inactivation, and a higher inhibition efficiency than the NSD1 inhibitor exhibited for its target.

Irreversible inactivation by **72** was also confirmed by dialysis (Figure 54). Pol β was inactivated following treatment with **72** (750 nM) and did not regain activity after dialysis. It is unclear why Pol β maintained 25% strand displacement activity before and after dialysis upon being exposed to 750 nM **72**. In several other experiments (Figures 51, 55, 56, and 57), a lower concentration of **72** (500 nM) at the same preincubation time (20 min) resulted in greater inactivation (i.e. <10% enzyme activity retained) of Pol β.

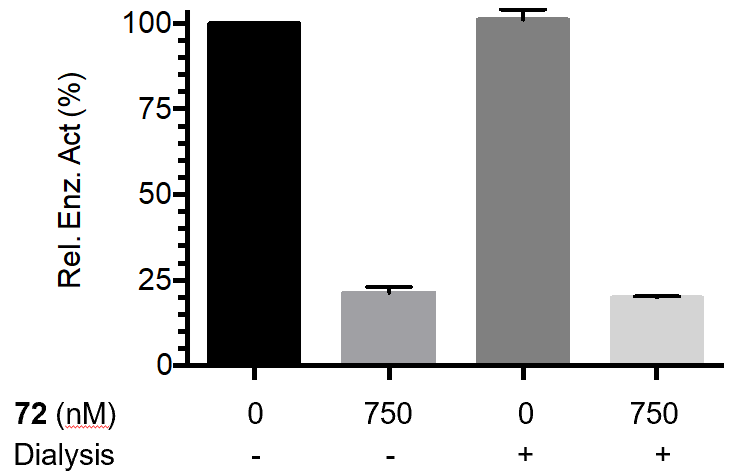


Figure 54. Strand displacement activity in response to dialysis of Pol β with **72**. Pol β (50 nM) was preincubated with or without **72** (750 nM) for 20 min. An aliquot of preincubation mixture was diluted 10-fold and mixed with 30 nM DNA substrate **B** and 300 μM dTTP. Another aliquot of the preincubation mixture (180 μL) was injected into a 3.5 K, 100-500 μL dialysis cassette and dialyzed against 1X reaction buffer (2 x 1 L, 50 mM HEPES, 50 mM KCl, 5 mM $MgCl_2$, 200 mM EDTA, 0.01% Tween 20, pH 7.5) for 24 h at 4 ºC. After dialysis, the strand displacement kinetics were analyzed following previously stated conditions.

The selectivity of **72** was characterized by testing it against other polymerases, including model replicative polymerase, Klenow exo$^-$, eukaryotic DNA repair polymerases, Pol θ and η, and the back-up BER polymerase, Pol λ.[25,161–163,169] Different assays were used to quantify inhibition for each polymerase (See Section 3.2.4).

Selectivity experiments showed that inhibitor **72** exhibited a strong preference for Pol β over these four polymerases even at 20-fold higher concentration (10 μM) of the inhibitor (Figure 55). At 500 nM **72**, all other polymerases tested retain full activity, while Pol β activity was completely reduced. Klenow exo$^-$ and Pol θ were also unaffected at 10 μM **72.** At a higher concentration of **72** (10 μM), the strongest inhibition occurs with Pol η and λ and ranges from 26-50% relative activity. The selectivity for Pol β over Pol λ is especially significant because they exhibit 32% sequence homology and previous Pol β inhibitors have also been potent against Pol

λ (i.e. **4** and HKL, Figure 20).[25,27,98–100] This degree of preference for Pol β has never been observed before in a potent (nM) inhibitor, which makes this result very exciting!

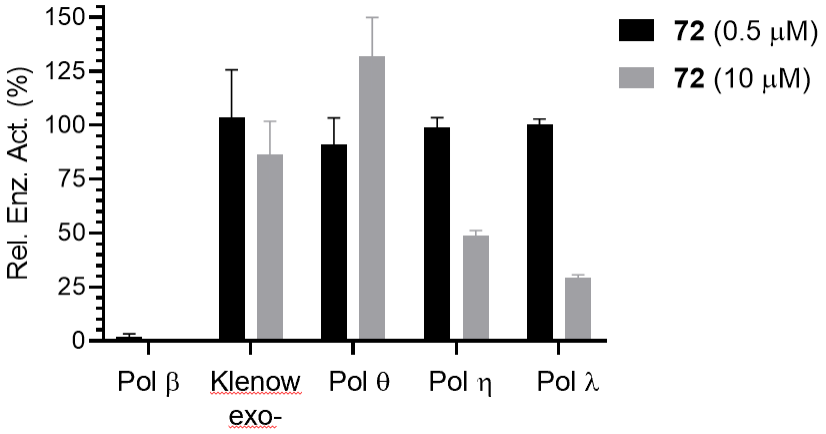


Figure 55. Selectivity of **72** against polymerase activity in model and mammalian polymerases. Polymerase (50 nM) was preincubated with or without **72** (500 nM, 10 μM) for 20 min. For fluorescence-based assays (Pol β, Klenow exo-, and Pol θ), an aliquot of preincubation mixture was diluted 10-fold and mixed with 30 nM DNA substrate and 300 μM dTTP (for Pol β and Klenow exo-) or 500 nM DNA and 400 μM dNTP (for Pol θ). For gel-based assays (Pol η and λ, an aliquot of preincubation mixture was diluted 10-fold and mixed with 125 nM DNA (Pol η, **D**; Pol λ, **E**) and 500 μM dNTPs at 37 ºC (for Pol η) or 10 nM DNA and 500 μM dNTPs at 37 ºC (for Pol λ), stopped at various time points by quenching with an equal volume of 95% formamide buffer, and analyzed by gel electrophoresis. Data are the average ± SEM for three experiments.

3.4.2.1 In Vitro biochemical and biophysical characterization of inhibitor 72

After more characterization, it was determined that the inhibition by **72** was dependent on pH and the concentration of thiols (Figures 56 and 57). Interestingly, Pol β activity at pH 6.9 and 8.0 was roughly half that at pH 7.4. While this was reproducible in our hands, it conflicted with other literature. Multiple reports suggest Pol β activity was enhanced at pH 8.0.[170,171] Compound **72** was most effective at physiological pH (7.4). Inactivation still occurred at slightly more acidic conditions (pH 6.9) but was completely disrupted at pH 8.0 (Figure 56). The pH effect was also observed after dialysis. When Pol β was preincubated at pH 7.4 with or without **72** (500 nM), Pol β activity was completely inactivated by **72** (denoted pH 7.4, checkered, Figure 56). When those

samples were dialyzed at pH 8.0 for 24 h and assayed for strand displacement activity at pH 7.4, there was no difference in activity between Pol β incubated with **72** (denoted pH 7.4*, checkered, Figure 56) and Pol β alone (denoted pH 7.4*, solid, Figure 56). The recovery of activity when comparing Pol β treated with **72** at pH 7.4 before (pH 7.4, checkered) and after (pH 7.4*, checkered) dialysis at pH 8.0 suggested that the inhibition was not only less effective at higher pH but was also reversible.

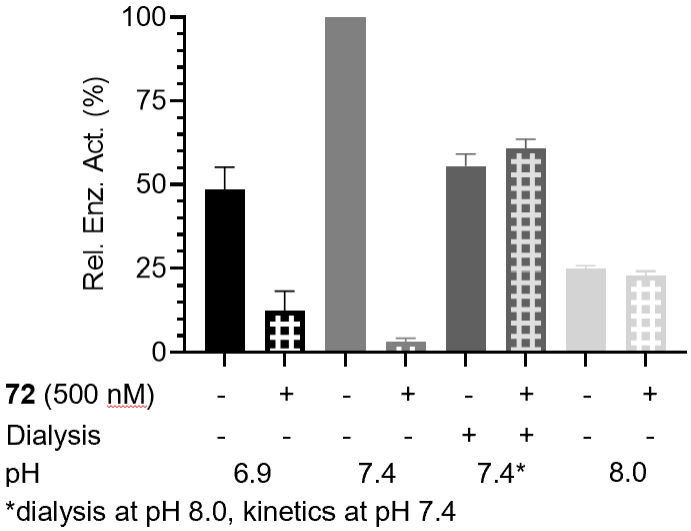


Figure 56. pH dependence of Pol β inactivation by **72**. Pol β (50 nM) was preincubated with (solid) or without (checkered) **72** (500 nM) for 20 min. An aliquot of preincubation mixture was diluted 10-fold and mixed with 30 nM DNA substrate **B** and 300 μM dTTP at various pH conditions (listed).

In addition, compound **72** was slightly less effective in the presence of thiol ([GSH] = 1, 5 mM) (Figure 57). Thiols decreased the efficacy of **72** in a concentration-dependent manner. This effect is common with electrophilic covalent inhibitors.[172,173] Thiols are strong nucleophiles that tend to decrease the effect of an electrophilic inhibitor by acting as a scavenger and decreasing the local concentration of active inhibitor available to inhibit the target enzyme (Scheme 30).

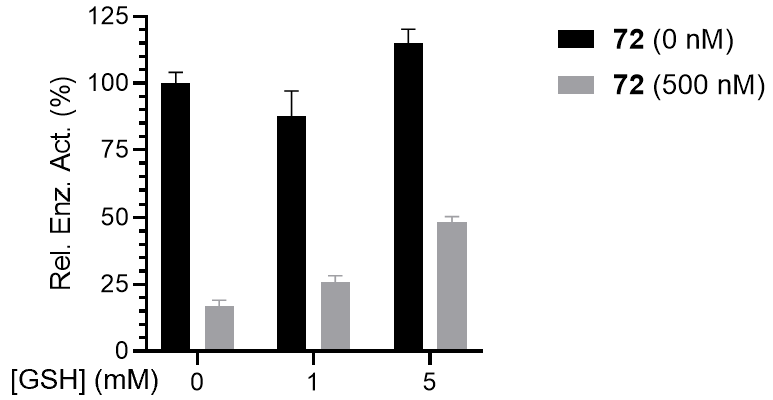
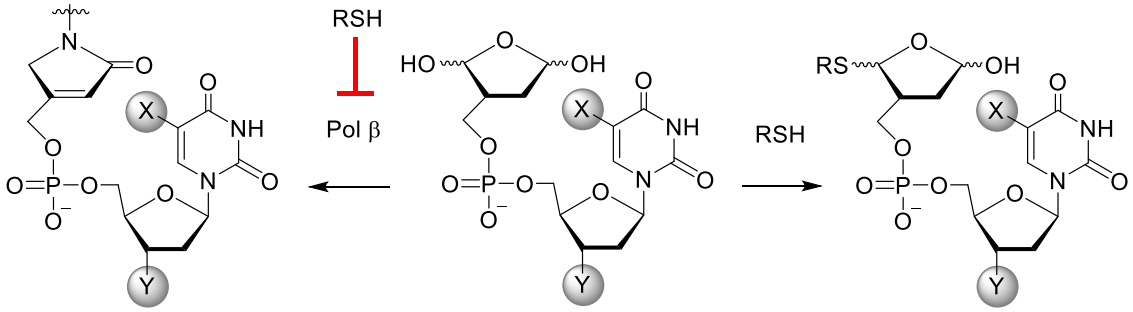


Figure 57. Thiols decrease the efficacy of **72**. Pol β (50 nM) was preincubated with or without **72** (500 nM) and GSH (0, 1, or 5 mM) for 20 min. An aliquot of preincubation mixture was diluted 10-fold and mixed with 30 nM DNA substrate **B** and 300 μM dTTP.

Scheme 30. Thiols react with electrophilic species.



We turned to mass spectrometry to determine the Pol β residue(s) modified by **72.** Covalent modification was detected on K113 and K234 (Figure 58, Fragments 11 and 12). The modification observed was a result of reaction between a Lys residue and the DOB fragment followed by loss of a water molecule (Figure 58, right). Fragment 11 contained a single internal lysine residue (K113). The role of K113 is unknown but lies near the DNA binding region (Figure 59). Fragment 12 contained 4 internal lysine residues (i.e. K230, K234, K244, and K248) and required MS/MS fragmentation to identify the lysine that was modified by compound **72.** Further analysis of the MS data identified an unmodified fragment containing K230 (b_{13}), an unmodified fragment containing K248 (y_9), and an unmodified fragment containing K244 and K248 (y_{18}). A fragment

with a single modification was observed that contained K230 and K234 (b_{21}*) and another fragment containing a single modification included K234, K244, and K248 (y_{21}*). Altogether, this suggested that of fragment 12, K234 was the modified lysine. K234 is known to be crucial for DNA binding interactions (Section 2.3.1.1.2).[115–117]

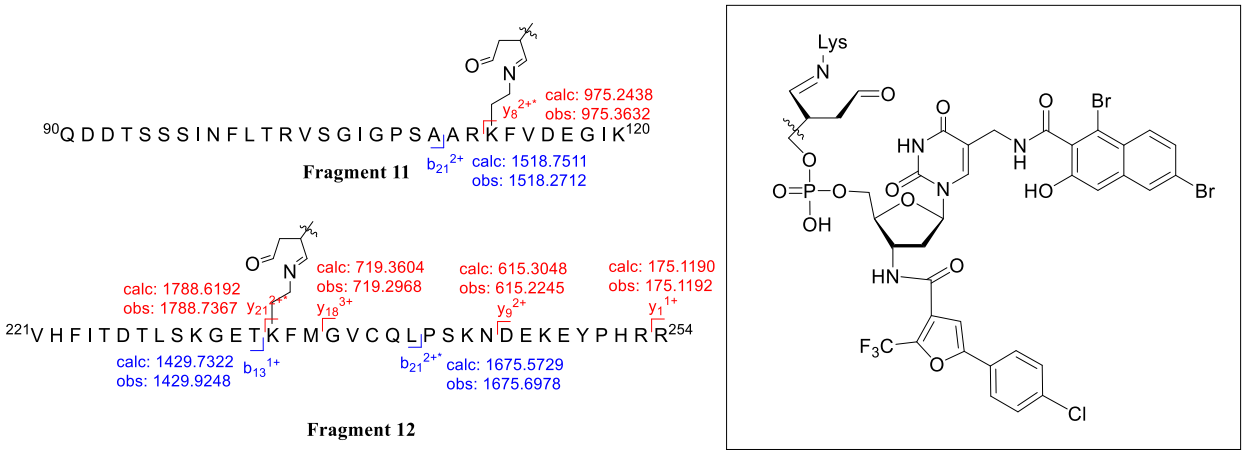


Figure 58. MS analysis of trypsin digestion of Pol β incubated with **72**. Pol β (1 μM) and **72** (300 nM) were incubated in 1X reaction buffer (50 mM HEPES, pH 7.2, 5 mM $MgCl_2$, 2 mM DTT) at 25 °C for 30 min. The mixture was concentrated with a 3 K Amicon filter and washed 3-4 times with 1X reaction buffer. Digestion buffer (25 μL), 500 mM Tris HCl, pH 7.0), 10X trypsin (25 μL, 400 μM), and H_2O (175 μL) was added. The digestion sample (2 μM Pol β and 40 μM trypsin) was incubated at 37 °C for 24 h. A portion (10 μL) of the digestion mixture was spun down and analyzed by UPLC-MS/MS using an ACQUITY UPLC HSS T3 Column (100 A, 1.8 μm, 2.1 mm x 100 mm). The flow rate was 0.3 mL/min running a gradient of 85:5:10 water: acetonitrile: 1% formic acid to 50:40:10 water: acetonitrile: 1% formic acid over 35 min. Analysis was conducted using BioPharmaLynx with tolerance set to 30 ppm and allowing for 4 missed cleavages.

Modification near the DNA binding regions led us to investigate whether inhibitor **72** impacted Pol β's ability to bind DNA (Figure 59). We turned to fluorescence anisotropy to measure DNA binding in the presence or absence of **72** (Figure 60). It is evident, based on fluorescence anisotropy, that inhibitor **72** interfered with DNA binding, especially at lower concentrations of Pol β. This effect was less prominent under high concentrations of enzyme because Pol β (200 nM) was in such large excess over the DNA (250 pM) that even if 99% of the enzyme was

inactivated, the remaining 1% of active enzyme was still in excess over the DNA. Altogether, MS analysis and fluorescence anisotropy support the hypothesis that inhibitor **72** binds to the polymerase domain and hindered DNA binding.

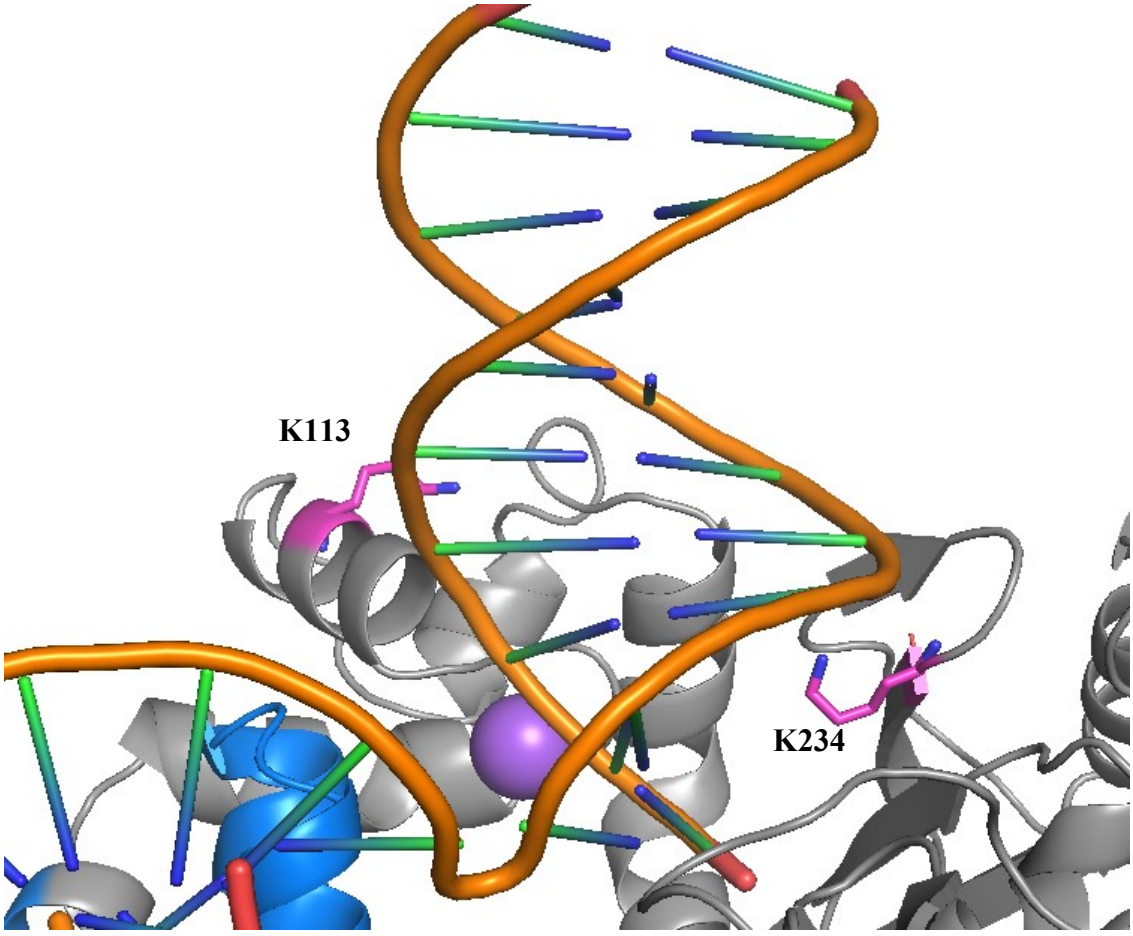


Figure 59. Crystal structure of WT Pol β: Covalently modified Lys residues after incubation with **72** (in pink) are in the DNA binding pocket.

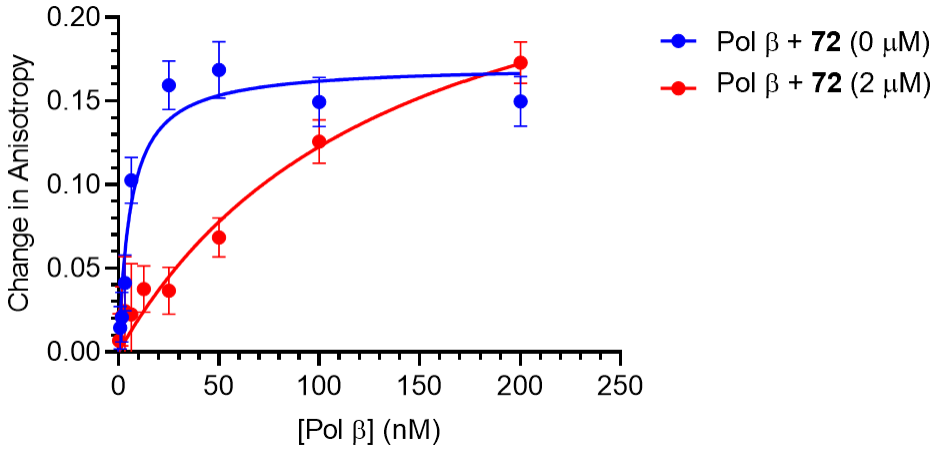
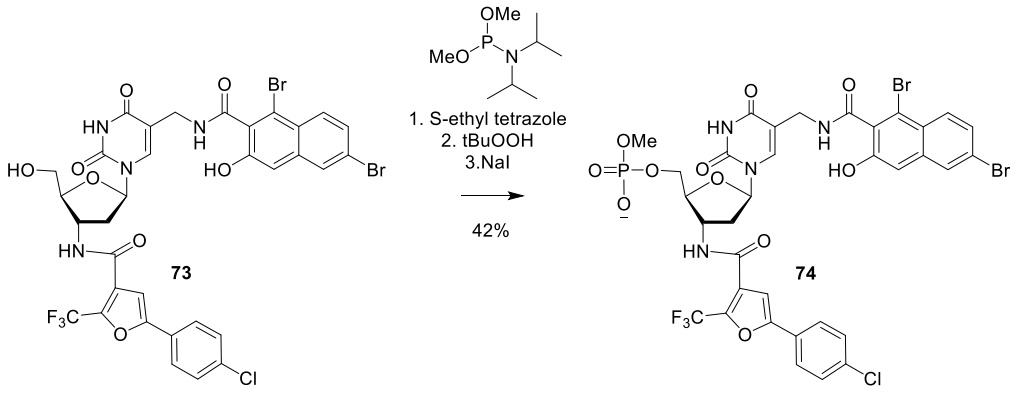


Figure 60. Fluorescence anisotropy experiments with **72**. Dichloro-diphenyl-fluorescein-labeled DNA (250 pM) was incubated with Pol β (various concentrations) in reaction buffer (50 mM HEPES, pH 7.5, 20 mM KCl, 1 mM EDTA, and 1 mM β-mercaptoethanol) in the presence of absence of inhibitor **72** (2 μM). Samples containing various concentrations of polymerase were prepared by serial dilution with 1X reaction buffer containing 250 pM DNA **F** and 2 μM **72** (to keep the concentration of DNA and **72** unchanged). The samples were incubated at 25 °C for 1 h and fluorescence anisotropy (A) was measured using a portion (125 μL) of each sample with PMT voltage 800 mV, 8 nm slit width, and 535 nm excitation and 556 nm emission.

3.4.2.2 Non-covalent analog of inhibitor 72

To test the significance of the DOB moiety, a noncovalent inhibitor analogue that resembled inhibitor **72** but instead of the DOB moiety bore a methyl ester was synthesized (Scheme 31). The 5'-hydroxyl nucleoside (**73**) containing both aryl groups was phosphorylated with a commercially available dimethyl phosphoramidite, oxidized, and one of the methyl groups was cleaved to yield the noncovalent inhibitor analogue (**74**).

Scheme 31. Synthesis of noncovalent inhibitor **74**.



The inhibition of Pol β's strand displacement activity by **74** was about 15-fold less potent than the inhibitor containing DOB (**72**) (Figure 61). We hypothesized that compound **74** would bind in the same area of the enzyme as **72** because of the similarity in structure despite compound **74**'s inability to covalently modify the enzyme. We used fluorescence anisotropy to determine if titration of **70** disrupted Pol β's ability to bind DNA. Based on competition fluorescence anisotropy experiments, the noncovalent inhibitor (**74**) also interfered with DNA binding, albeit at a much higher concentration (> 15 μM) than **72** (2 μM) (Figure 54). While the data confirm that covalent modification by the DOB moiety enhanced Pol β inhibition, there was a discrepancy between the disruption of DNA binding interactions and the observed effect on enzyme activity for the noncovalent inhibitor (**70**). Based on fluorescence anisotropy data, < 5 μM **74** did not affect DNA binding but according to the strand displacement kinetics, 500 nM **74** decreased Pol β activity by ~25%.

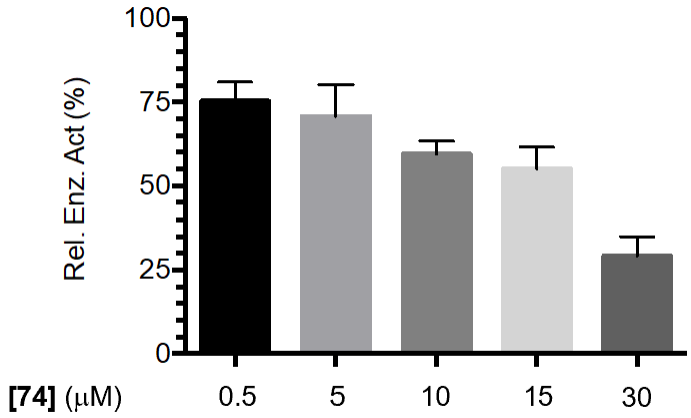


Figure 61. Inactivation of strand displacement activity by **74** in WT Pol β. Polymerase β (50 nM) was preincubated with or without **74** (various concentrations) for 20 min. An aliquot of preincubation mixture was diluted 10-fold and mixed with 30 nM DNA substrate **B** and 300 μM dTTP.

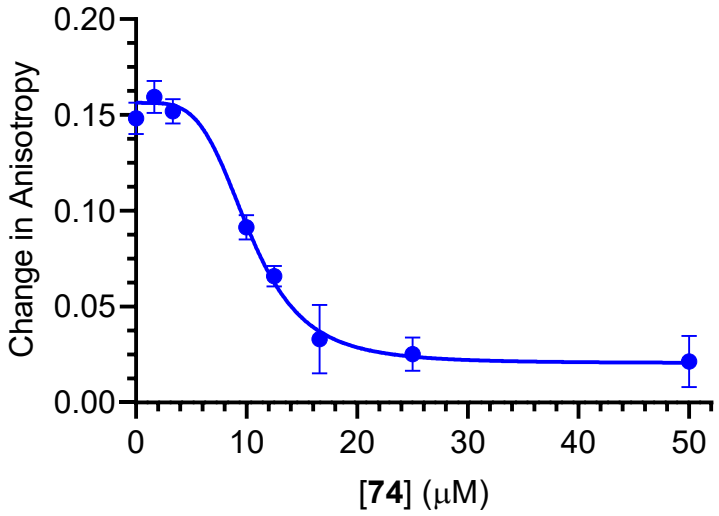
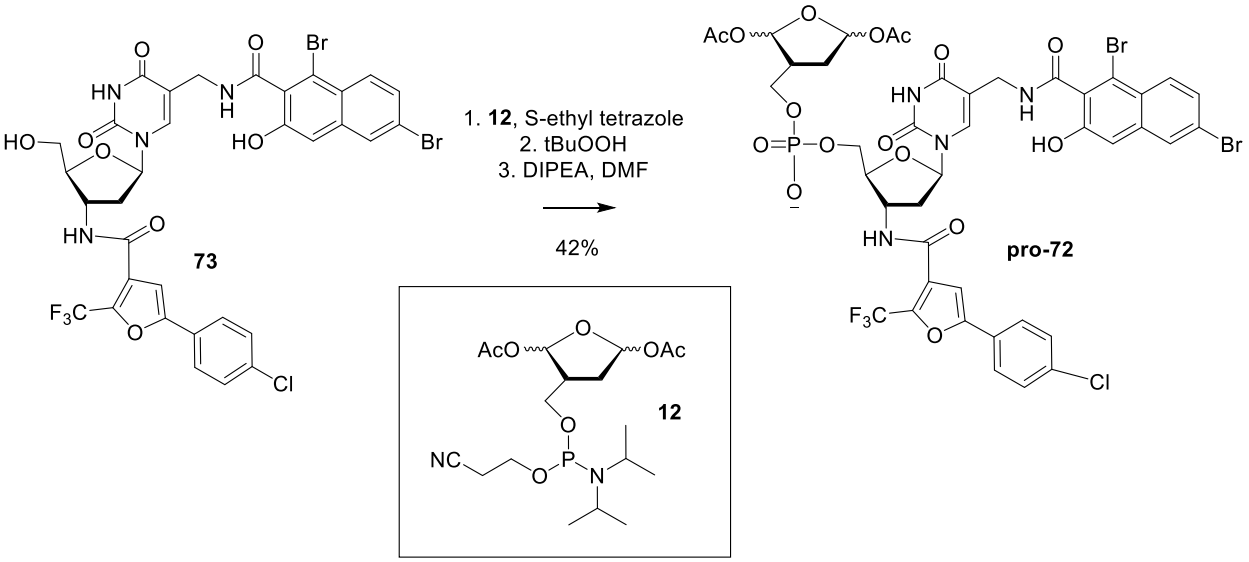


Figure 62. Competition fluorescence anisotropy with **74** in WT Pol β. Dichloro-diphenyl-fluorescein-labeled DNA **F** (250 pM) was incubated with Pol β (25 nM) in reaction buffer (50 mM HEPES, pH 7.5, 20 mM KCl, 1 mM EDTA, and 1 mM β-mercaptoethanol) with various concentrations of **74** (0-50 μM). After measuring anisotropy of initial sample containing **74** (1.6 μM), the next sample was prepared by titrating additional **74**, allowing mixture to equilibrate for 10 min and measuring anisotropy.

3.4.2.3 The effect of pro-72 in cells

The second-generation hit compound was also evaluated in cells using a pro-inhibitor (**pro-72**). The pro-inhibitor was synthesized by coupling the 5'-OH nucleoside (**73**) bearing both aryl groups with previously synthesized phosphoramidite (**12**) (Scheme 32).

Scheme 32. Synthesis of pro-inhibitor (**pro-72**).



The toxicity of **pro-72** was evaluated in HeLa cells using cell viability and clonogenic assays. HeLa cells were treated with **pro-72** (0, 5, 25, 50, 100 μM) for 1 or 2 h. Cell viability assays indicated that **pro-72** killed less than 10% of the cells at these concentrations (Figure 63). A clonogenic assay of HeLa cells treated with **pro-72** showed no significant cell death at low concentrations (5, 25 μM) regardless of the treatment time (Figure 64). Significant toxicity occurred at 100 μM **pro-72** treatment for 2 h (Figure 64).

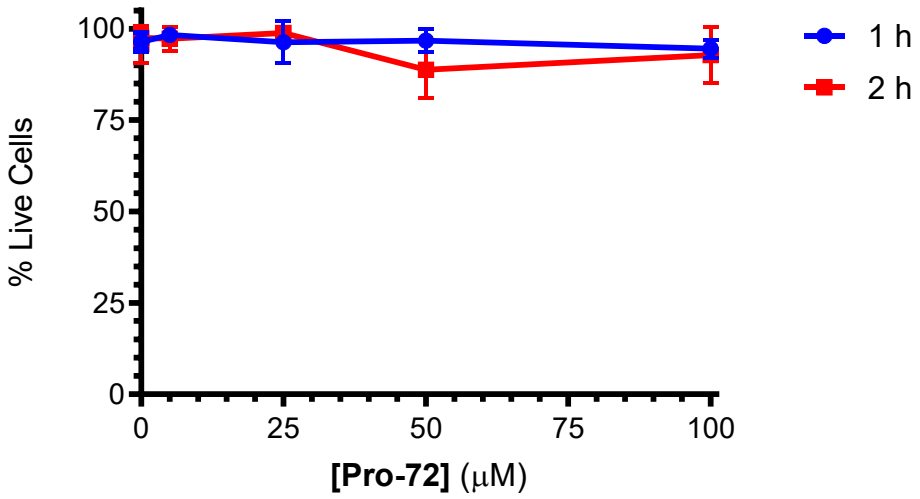


Figure 63. Cell viability of HeLa cells in response to **pro-72**.

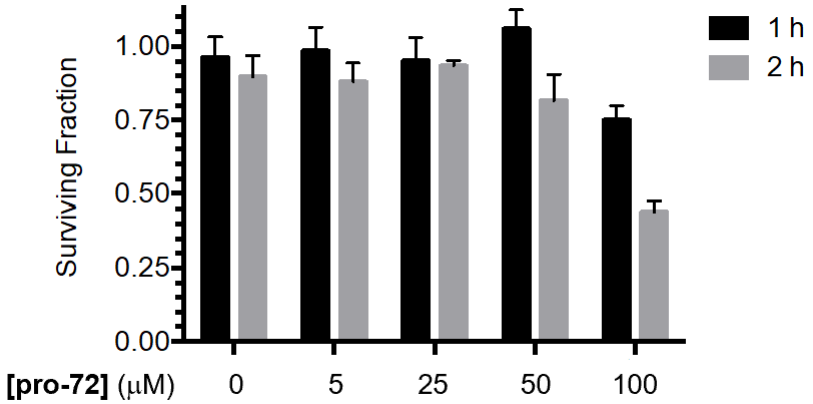


Figure 64. Clonogenic assays show low toxicity of **pro-72** in HeLa cells. HeLa cells were treated with **pro-72** (various concentrations) for 1 or 2 h. After treatment, media containing **pro-72** was removed, cells were harvested, counted, and seeded to allow for colony growth. Colonies were visualized by crystal violet.

Nontoxic concentrations of pro-inhibitor were ideal for determining synergism with a DNA damaging agent. Synergism implies that the cell death resulting from combination treatment (AB) is greater than the additive cell death observed from individual treatments (A + B) (equation 6). Therefore, if an individual treatment was too toxic to the point where the additive effect was close

to or greater that 100% cell death, no synergism can be identified. Therefore, we used 5 or 25 μM **pro-72** in all clonogenic assays in the presence or absence of MMS or BLM.

$$AB/(A+B) > 1 \quad (6)$$

A synergistic effect with MMS and BLM was tested for via a clonogenic assay at nontoxic concentrations of **pro-72** (5, 25 μM) (Figure 65). There was a synergistic effect between **pro-72** (5 μM) and MMS. Combination treatment with MMS (0.2 mM) and **pro-72** (5 μM) for 1 h was 2.5 times more effective at killing cells than the same treatment without **pro-72**. A longer treatment (2 h) of MMS killed 58% cells while the combination treatment resulted in 97% cells death, a 1.7-fold enhancement. Curiously, for a 1 h treatment this synergistic effect was not enhanced by 25 μM **pro-72** (light blue, Figure 65). Synergism using a longer treatment would likely be difficult to observe due to the extensive cell death already observed at 5 μM **pro-72.**

The synergistic effect between **pro-72** and BLM was slightly greater than with MMS (Figure 65). A 1 h treatment of **pro-72** (5 μM) and BLM (2 μM) enhanced the cell killing effect by 2.8-fold compared to the single treatment of BLM at the same concentration. The 2 h treatment of BLM was too toxic to observe a synergistic effect with **pro-72**. In the same way that nontoxic conditions of **pro-72** were ideal, nontoxic (or little toxicity) conditions of DNA damaging agents were preferable for measuring synergism. Altogether, these data indicate our Pol β pro-inhibitor enhanced the efficacy of DNA damaging agents (i.e. MMS and BLM) in HeLa cells.

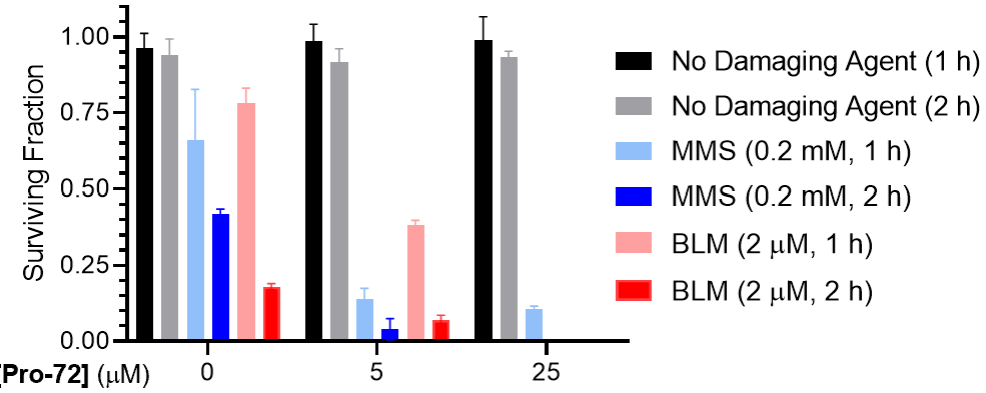


Figure 65. Synergistic effect of **pro-72** with DNA damaging agents on cell death in HeLa cells. HeLa cells were treated with MMS (0.2 mM) or BLM (2 μM) in the presence or absence of **pro-72** (various concentrations) for 1 or 2 h. After treatment, media containing **pro-72** was removed, cells were harvested, counted, and seeded to allow for colony growth. Colonies were visualized by crystal violet.

The effectiveness of **pro-72** was also tested in a breast cancer cell line, MCF-7 cells. We observed that HeLa cells were more robust than MCF-7 and we were unsure how the pro-inhibitor would fare in a more sensitive cell line. It was clear that **pro-72** was generally more toxic in MCF-7 (Figure 66) than in HeLa (Figure 64) cells. While a 1 h treatment of **pro-72** (5 μM) resulted in 10% cell death in MCF-7 (Figure 66) or HeLa cells (Figure 64), a 2 h treatment of **pro-72** (5 μM) in MCF-7 cells increased cell death to 39% (Figure 66). Interestingly, we do not observe a significant change in cell death between 5 and 25 μM **pro-72** regardless of the treatment time. A similar effect was observed in HeLa cells (Figure 64).

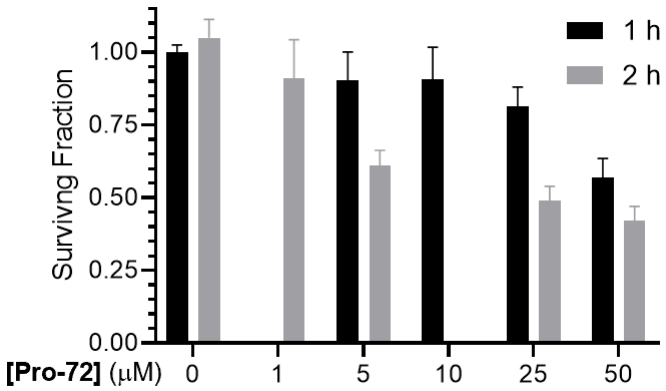


Figure 66. Toxicity of **pro-72** in MCF-7 cells as determined via clonogenic assay. MCF-7 cells were treated with **pro-72** (various concentrations) for 1 or 2 h. After treatment, media containing **pro-72** was removed, cells were harvested, counted, and seeded to allow for colony growth. Colonies were visualized by crystal violet.

It is worth mentioning that biological replicates of cell experiments are more variable than in vitro experiments.[174,175] For example, the same treatment of **pro-72 (**5 μM; 2 h) in MCF-7 cells resulted in 45% cell death in one set of experiments (Figure 66) and 25% cell death in another set of experiments (Figure 67). This is important to keep in mind when analyzing data.

To compare the activity of **pro-72** in HeLa and MCF-7 cells, we used clonogenic assays to measure a synergistic effect between **pro-72** and DNA damaging agents in MCF-7 cells. When MCF-7 cells were treated with MMS or BLM in the absence or presence of **pro-72**, a synergistic effect at killing cells was observed (Figure 67). The synergism between MMS and **pro-72** was greater than that of BLM and **pro-72**. Treatment with MMS (0.2 mM) in the presence of **pro-72** (5 μM) was 1.5 times more effective at killing cells that were treated with MMS (0.2 mM) for 1 h. However, a 2 h treatment of the DNA damaging agent was too harsh and resulted in too much cell death to accurately measure a synergistic effect. These experiments must be repeated under less toxic conditions to establish a synergistic effect upon longer treatments. Interestingly, a synergistic effect was not observed between BLM (2 μM) and **pro-72** (5 μM). Considering that treating MCF-

7 cells with BLM (2 μM) for 1 h killed 60% of the cells and **pro-72** killed 30%, this treatment was likely too harsh on the cells and it is possible a synergistic effect may be observed under less toxic conditions.

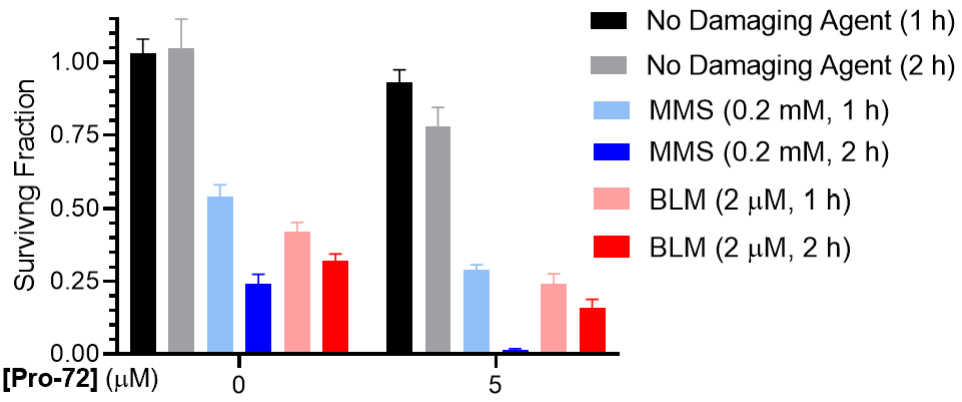


Figure 67. Synergistic effect of **pro-72** with DNA damaging agents on cell death in MCF-7 cells. MCF-7 cells were treated with MMS (0.2 mM) or BLM (2 μM) in the presence or absence of **pro-72** (various concentrations) for 1 or 2 h. After treatment, media containing **pro-72** was removed, cells were harvested, counted, and seeded to allow for colony growth. Colonies were visualized by crystal violet.

Altogether, **pro-72** was more toxic in MCF-7 cells than in HeLa cells. In addition, synergism between **pro-72** and DNA damaging agents in MCF-7 cells (Figure 67) was less impressive than synergism observed in HeLa cells (Figure 65) but still observed.

3.5 Identification of Pol β mutant inhibitor(s)

A potent inhibitor that is selective for a mutant form of Pol β associated with a specific cancer cell is highly desirable (Section 2.5.1). Improving the preference for a tumor over a healthy cell has the potential to enhance the agent's therapeutic index (TI, Section 1). Therefore, another aspect of this research was identifying and characterizing an inhibitor of a mutant form of Pol β.

3.5.1 Mutants Screened

Mutant forms of Pol β have been identified in patient tumors.[13] The degree to which each identified mutant has been characterized (i.e. relative activity) differs. Three plasmids containing active lyase mutants (i.e. N24K, P63L, and T79I) were acquired and expressed (Figure 68). At the time that we obtained the expression plamids for these mutants, we were focused on lyase mutants because our inhibitors were based on an oxidized abasic lesion (i.e. DOB), which covalently modified Lys72 of the lyase domain. We had requested the mutants prior to determining that inhibitor **2** or **72** targeted the polymerase domain (Section 3.3.2 and Section 3.4).[29]

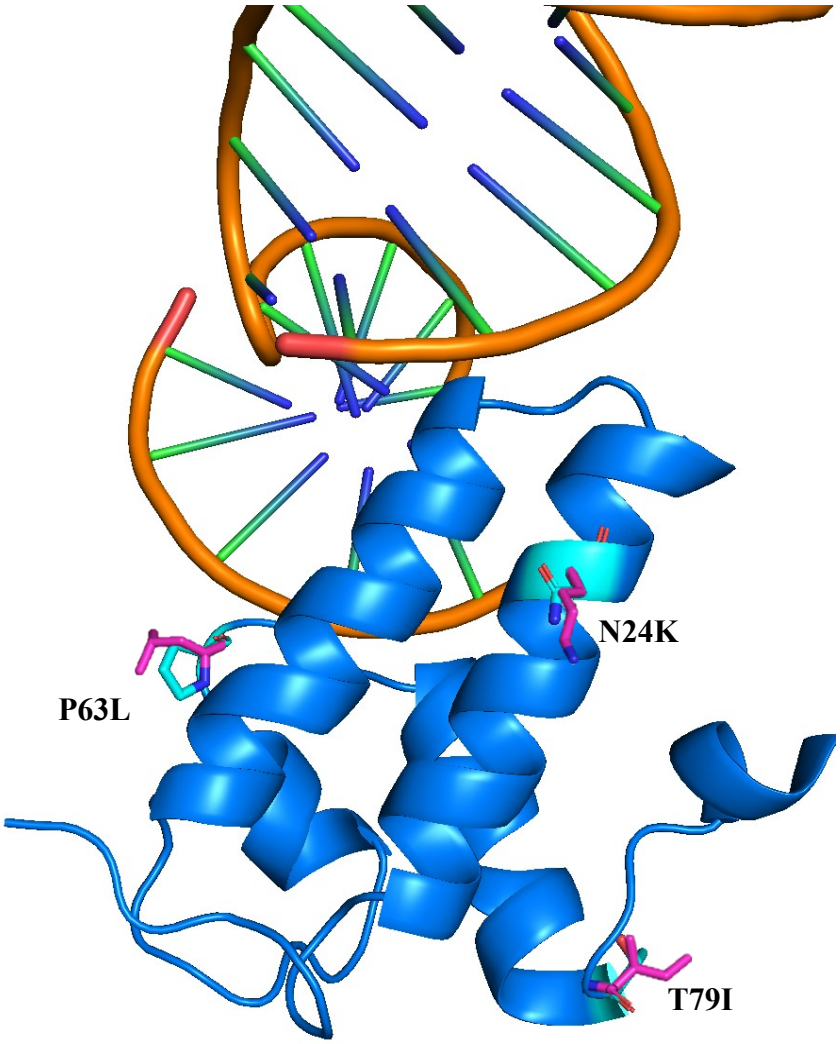
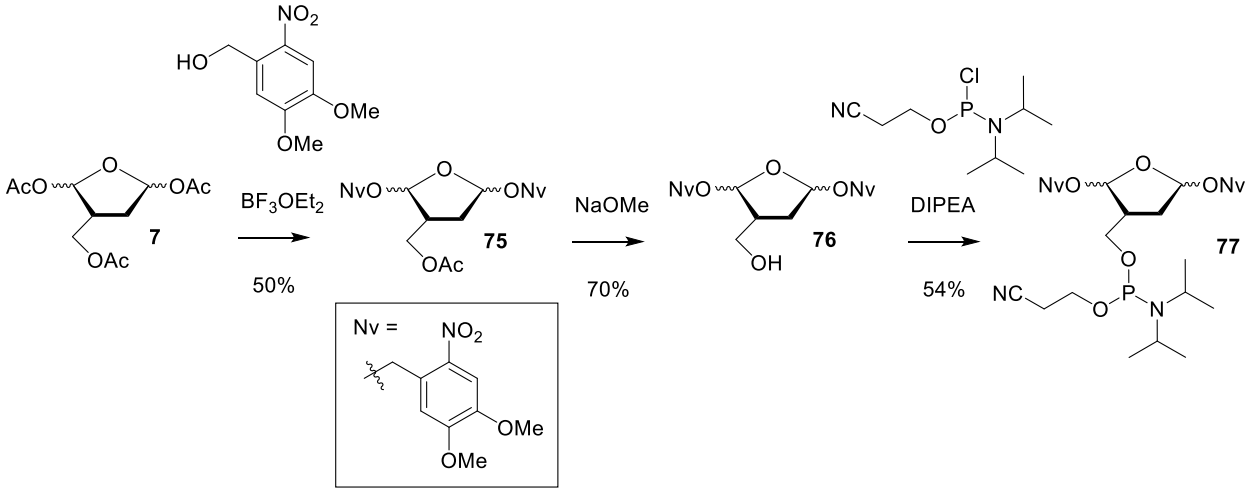


Figure 68. Crystal structure of the mutants (in pink) in the 8 kDa Pol β lyase domain (blue) bound to DNA (PDB: 1PBX).

3.5.2 First-generation mutant inhibitor: Functionalization at the 3'-position

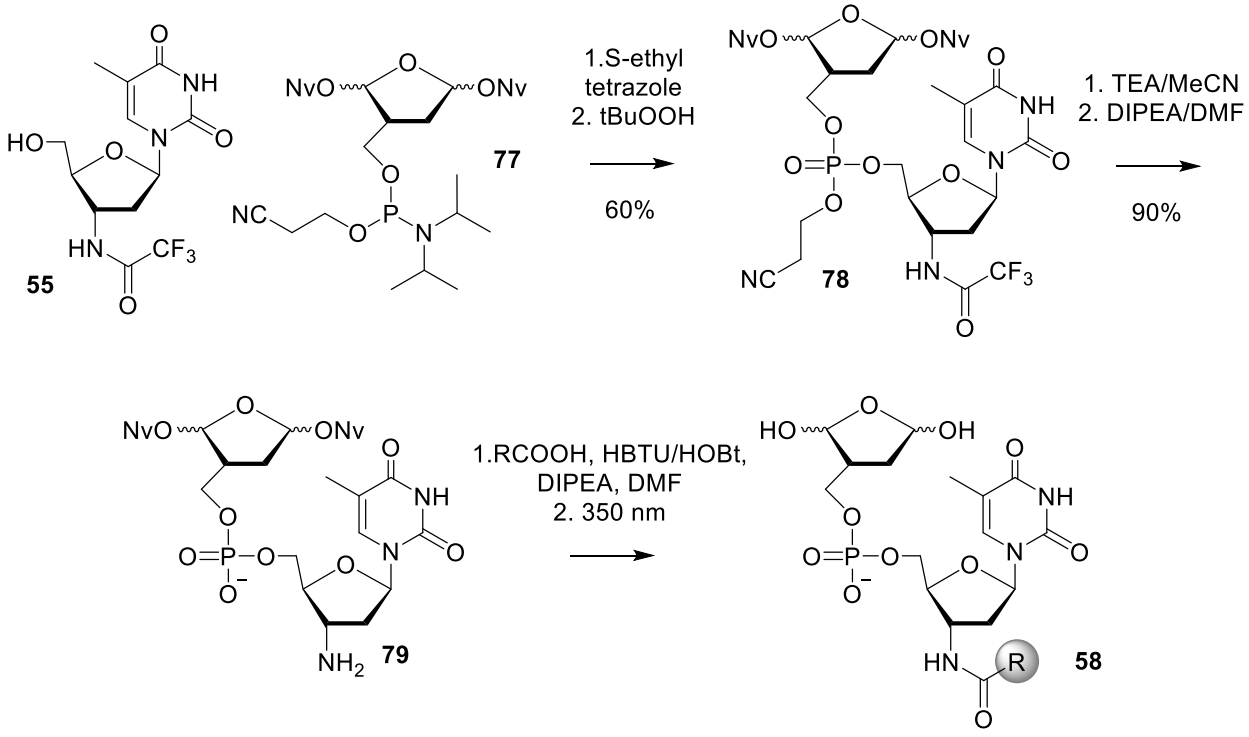
The first-generation library for Pol β mutants was the same as the first-generation library for WT Pol β. However, this library synthesis was altered to avoid the pentenoyl protecting groups on the DOB because the NBS deprotection of these groups was not trivial. More importantly, we determined that the deprotection conditions (i.e. NBS) were incompatible with several easily oxidized carboxylic acids (i.e. hydroxylated naphthoic acids). Instead, nitroveratryl (Nv) groups were used to protect the DOB (Scheme 33). Nv groups have been used previously in nucleic acid chemistry as photolabile precursors.[26,27,29,135] Prior to synthesizing a DOB phosphoramidite with Nv protecting groups, I confirmed that photochemical cleavage would be compatible with the carboxylic acids that were incompatible with NBS deprotection. Nv groups are typically removed by exposure to 365 nm light. Naphthoic acids absorb light at this wavelength but are unreactive under the photolysis conditions used. Therefore, we were confident that this approach would be compatible across the library of carboxylic acids. Previously synthesized **7** (Scheme 12) was transformed into Nv-protected **75** in the presence of a Lewis acid (i.e. BF_3 etherate) and the commercially available 6-nitroveratryl alcohol. Following deacetylation (**76**) and phosphitylation, the diastereomers of phosphoramidite **77** were separately isolated.

Scheme 33. Synthesis of Nv-protected DOB moiety.



To generate the Nv-protected scaffold molecule for the same library (**58**) previously synthesized, **55** was coupled to **77** to yield phosphate triester **78** (Scheme 34). After phosphate and amine deprotection, the Nv-protected scaffold (**79**; 30 mg total, 100 nmol per carboxylic acid) was coupled to a library of carboxylic acids (~375 compounds) and deprotected in a mixture of 1:1 MeCN/H_2O by exposure to 365 nm light for 15 min. The deprotection of the Nv group was confirmed by MS and HPLC. Furthermore, the deprotection conditions were flexible. Successful deprotection was observed after as short at 10 min and as long as 30 min exposure to light with minimal decomposition.

Scheme 34. Synthesis of Nv-protected scaffold **79** and library **58**.



This library was screened against the lyase mutants (N24K, P63L, and T79I), as well as WT Pol β using the fluorescence-based UBER lyase assay (Section 3.2.3, Figure 27) in order to identify any hit compounds selective for a mutant over the WT enzyme. This assay was important

because the use of the strand displacement assay for high-throughput screening may be ineffective for these purposes. All mutants tested contain the WT 31 kDa domain. Therefore, any selectivity identified in the strand displacement assay for the mutants would be based upon a distal modification and may be unreliable.

Each compound (25 μM) was preincubated with WT and mutant Pol β (100 nM) for 25 min, incubated with substrate **A** for 25 min (Table 1), then diluted 5-fold, and mixed with the UBER probe (Figure 27). The first screen identified several inhibitors, so all hit compounds were resynthesized and rescreened to rule out false positives.

The fluorescence-based lyase assay identified two hit compounds: **80** and **81** for Pol β N24K and T79I, respectively (Figure 69). Again, the presence of the naphthoic acid moiety has been commonly found in inhibitors of Pol β. While the UBER assay was an excellent screening method, it was not sensitive enough to obtain quantitative kinetic information about lyase inactivation. Therefore, the gel-based lyase assay was used to evaluate these candidates further (Section 3.2.2).

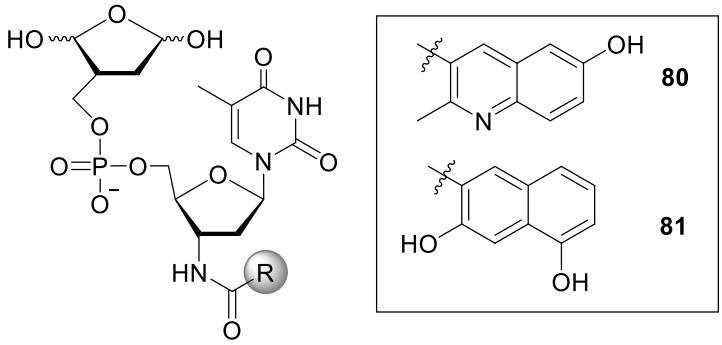
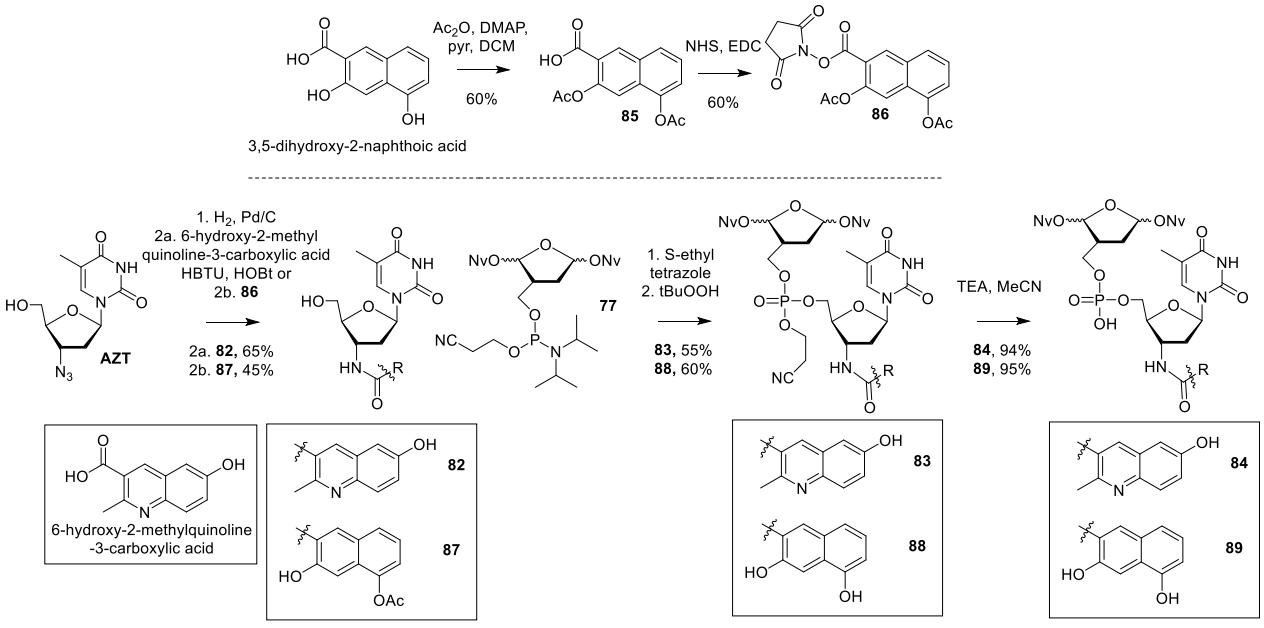


Figure 69. Identification of hit compounds **80** and **81** against for Pol β N24K and Pol β T79I, respectively, using the UBER lyase assay.

Compounds **80** and **81** were resynthesized and purified on a larger scale (5-10 mg) because we determined the crude compounds were contaminated by activating agents (i.e. HBTU, HOBt). We found that completing the amide coupling with the corresponding carboxylic acid prior to

phosphoramidite coupling resulted in more pure final compounds (Scheme 35). In the synthesis of **80**, AZT was reduced and coupled to 6-hydroxy-2-methylquinoline-3-carboxylic acid to yield **82**. Phosphoramidite coupling with **77** yielded phosphate triester **83.** Following deprotection, the precursor (**84**) to inhibitor **80** was isolated. In the synthesis of **81**, AZT was reduced and coupled to the NHS ester (**86**) of the Ac-protected hydroxylated naphthoic acid (**85**) to yield **87**. This intermediate (**87**) was coupled to phosphoramidite **77** to yield **88**. Compound **88** was deprotected to yield the precursor (**89**) to inhibitor **81**. Each precursor (**84** and **89**) was converted to the active inhibitor (**80** and **81**) by cleaving the photolabile Nv group upon exposure to light.

Scheme 35. Synthesis of first-generation mutant hits.



These hits were analyzed via the quantitative gel-based lyase assay (Section 3.2.2) and the qualitative UBER assay (Section 3.2.3). Regardless of the assay used, **80** was more potent against the N24K mutant than the WT enzyme (Figure 70). Inhibitor **80** reduced WT Pol β lyase activity

by about 30% in the gel-based lyase assay, while it had no effect on WT Pol β in the UBER assay. Meanwhile, inhibitor **80** reduced Pol β N24K lyase activity to about 35-40%, regardless of the assay used to quantify. By comparing the effect of inhibitor **80** in the separate assays, we observed a 35% difference in favor of Pol β N24K in the gel-base lyase assay versus a 60% difference in favor of Pol β N24K in the UBER assay. Therefore, we were wary that inhibitor **80** looked more selective for the N24K mutant against WT in the UBER assay, compared to the gel-based lyase assay. Across assays, **81** was less potent and less selective towards mutant T79I when comparing it to **80** with N24K.

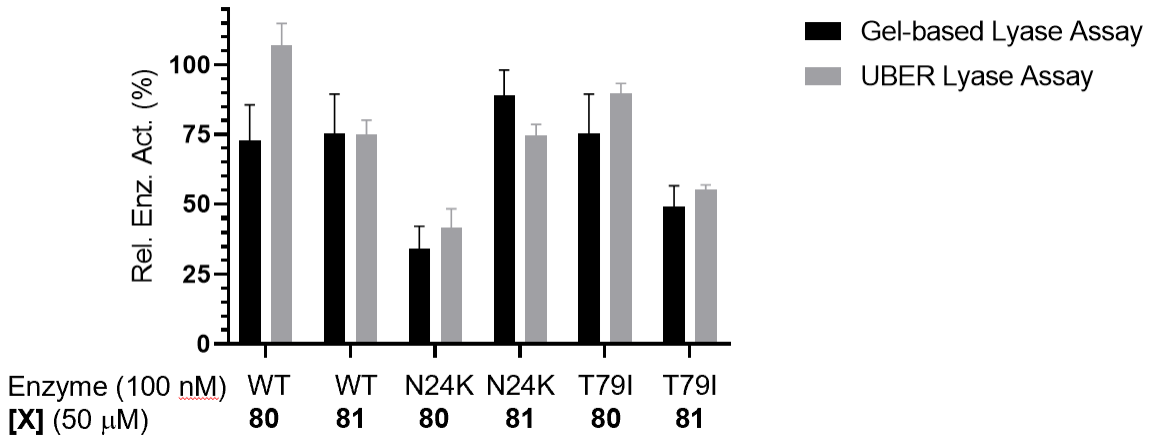


Figure 70. Gel-based and UBER lyase assay comparison in Pol β inactivation by hit compounds. Pol β or mutant (100 nM) was preincubated with hit compounds **80** or **81** (50 μM) for 20 min. The preincubation mixture was diluted 10-fold or 5-fold and mixed with DNA substrate **A** for the gel-based lyase assay or the UBER lyase assay, respectively. In the gel assay, aliquots were removed and frozen at various time points (0, 2, 4, 6, 8, 10 min), quenched with $NaBH_4$, and analyzed by PAGE. In the UBER assay, the enzyme/DNA mixture incubated for 25 min and mixed with the UBER probe for 4 h prior to scanning with a Clariostar wellplate reader.

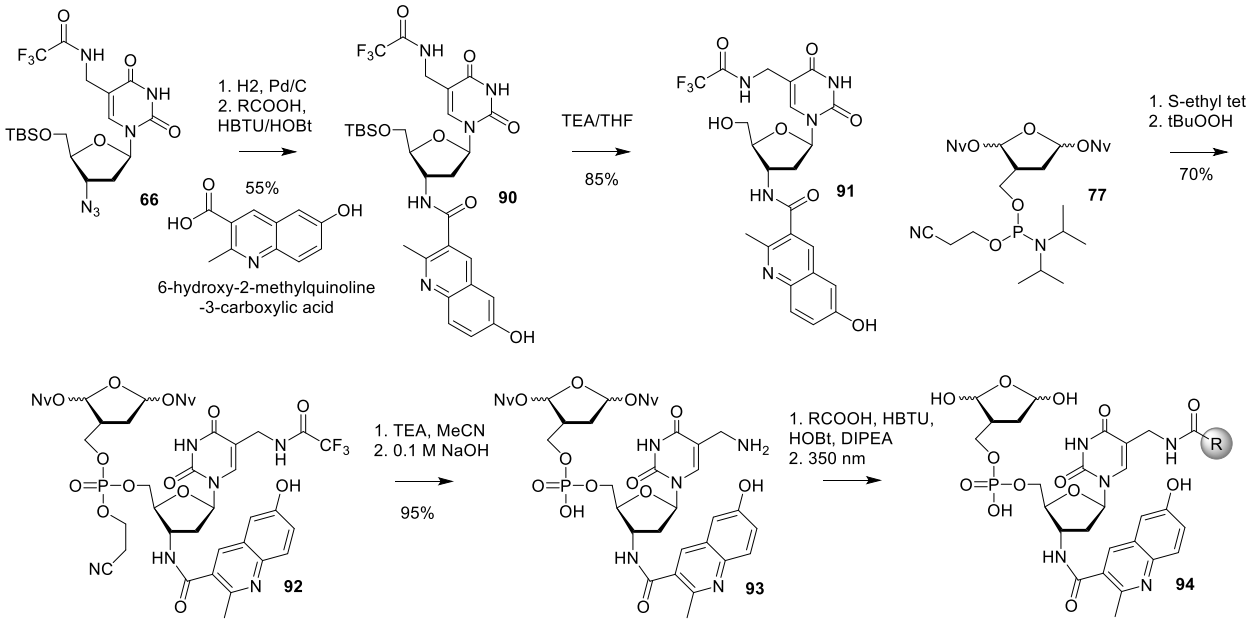
It was evident that these compounds were not very potent (~50 μM) but we pursued this further with the hope that potency and selectivity may improve in a second-generation inhibitor. Therefore, we synthesized a second-generation iteration of inhibitor **80** because it was more potent,

and slightly more selective for the N24K mutant. In addition to this, we thought the N24K may be a better candidate for targeting because the position of the mutant lies near the inside of the enzyme and towards the DNA (Figure 64). On the contrary, P63L and T79I are solvent-exposed and we expect it to be difficult to make specific contacts that would lead to inactivation.

3.5.3 Second-generation mutant inhibitor(s): Functionalization at the thymine 5-methyl position

The synthesis of the second-generation scaffold was straight forward and similar to the synthesis of the previous second-generation scaffold for WT Pol β (Scheme 28). Previously synthesized **66** was reduced and coupled to the carboxylic acid used in the synthesis of **80** to yield compound **90** (Scheme 36). After silyl deprotection, the 5'-hydroxylated quinoline (**91**) was coupled to the Nv-protected DOB phosphoramidite (**77**) to form **92**. Following deprotections, scaffold **93** was coupled to a library of 375 carboxylic acids (Appendix 1) and photochemically deprotected to make library **94**.

Scheme 36. Synthesis of second-generation library (**94**).



Library **94** (15 μM) was screened against WT Pol β and Pol β N24K (100 nM) with a 25 min preincubation period using the UBER assay (Section 3.2.3). We screened the library at 15 μM because the scaffold partially inactivated Pol β at 20 μM. Following the 25 min preincubation, substrate **A** was added to the enzyme/inhibitor solution and incubated for an additional 25 min (Table 1). The mixture was diluted 5-fold and mixed with the UBER probe (5 μM) (Figure 27). Similar to the first-generation screen, several inhibitors were identified, so all hit compounds were resynthesized and rescreened, along with the free carboxylic acid control experiments, to rule out false positives.

After three rounds of screening, 8 remaining compounds (**95-102**) showed partial inactivation of Pol β N24K in the UBER assay (Figure 67). However, the UBER assay was not used for quantitative analysis, so these compounds were analyzed via the gel-based lyase assay (Section 3.2.2).

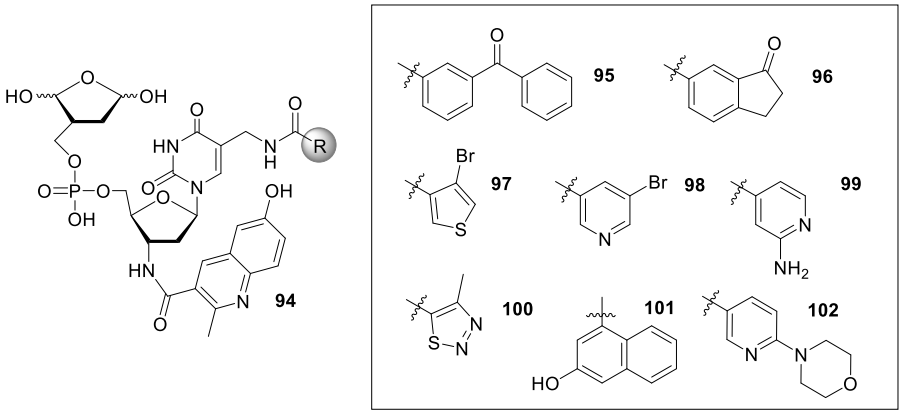


Figure 71. Hit compounds identified from UBER screen.

It was evident none of the hit compounds **95-102** did not selectively inhibit Pol β N24K over the WT Pol β based on the gel-based lyase assay (Figure 72). Furthermore, the inhibitors were not very potent at 15 μM (i.e. the greatest effect was roughly 50% enzyme activity). In summary, the hits identified were not impressive in terms of potency or selectivity for the mutant enzyme.

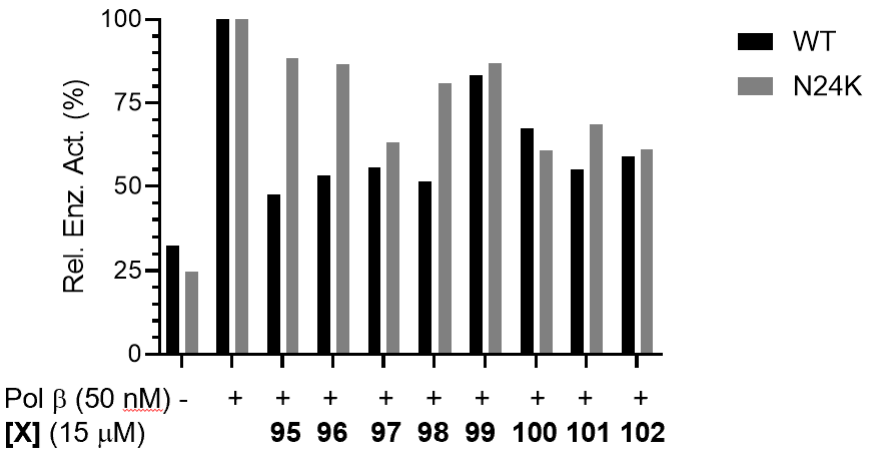


Figure 72. Gel-based lyase assay analysis of mutant hits. Pol β or mutant (100 nM) was preincubated with hit compounds (15 μM) for 20 min. The preincubation mixture was diluted 10-fold and mixed with DNA substrate **A**. Aliquots were removed and frozen at various time points (0, 2, 4, 6, 8, 10 min), quenched with $NaBH_4$, and analyzed by PAGE.

This led us to question why we did not observe an enhanced inhibitor in the second-generation library. We hypothesized that if the inhibitor(s) were binding in the polymerase domain, we would not observe any selectivity because the mutant and the WT have the same 31 kDa domain. We chose the most effective hits (**97**, **100**, and **102**) and tested them in the 8 kDa domain (Figure 73). It was very clear that the second-generation hits and the first-generation compound (**80**) had no effect on the lyase activity of the 8 kDa domain. This supported our hypothesis that our compounds may be binding in the polymerase domain.

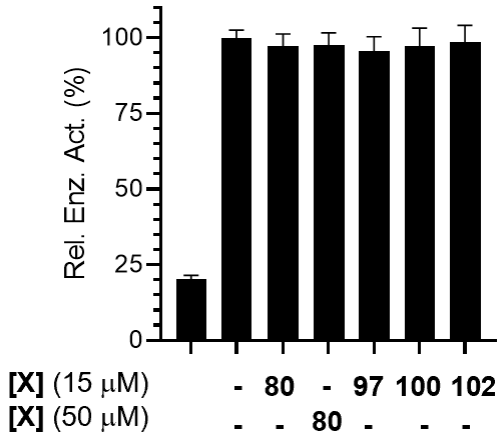


Figure 73. Mutant hits do not inactivate 8 kDa domain. 8 kDa Pol β (100 nM) was preincubated with hit compounds (15 or 50 μM) for 20 min. The preincubation mixture was diluted 10-fold and mixed with DNA substrate **A**. Aliquots were removed and frozen at various time points (0, 2, 4, 6, 8, 10 min), quenched with $NaBH_4$, and analyzed by PAGE.

To test this further, we analyzed our hits via the fluorescence-based strand displacement assay in the WT Pol β and the 31 kDa domain (Section 3.2.1). The second-generation hits (**97**, **100**, **102**) at 15 μM did not significantly inhibit the activity in the WT (Figure 74a) or the polymerase domain (Figure 74b), which was consistent with the gel-based lyase assay (Figure 72). A more significant decrease in strand displacement activity was observed for **80** (50 μM) with WT enzyme and the 31 kDa domain (dark blue, Figure 74). This was consistent with WT Pol β inactivation by **80** (50 μM) to about 35% relative activity, according to the gel-based lyase assay (Figure 70).

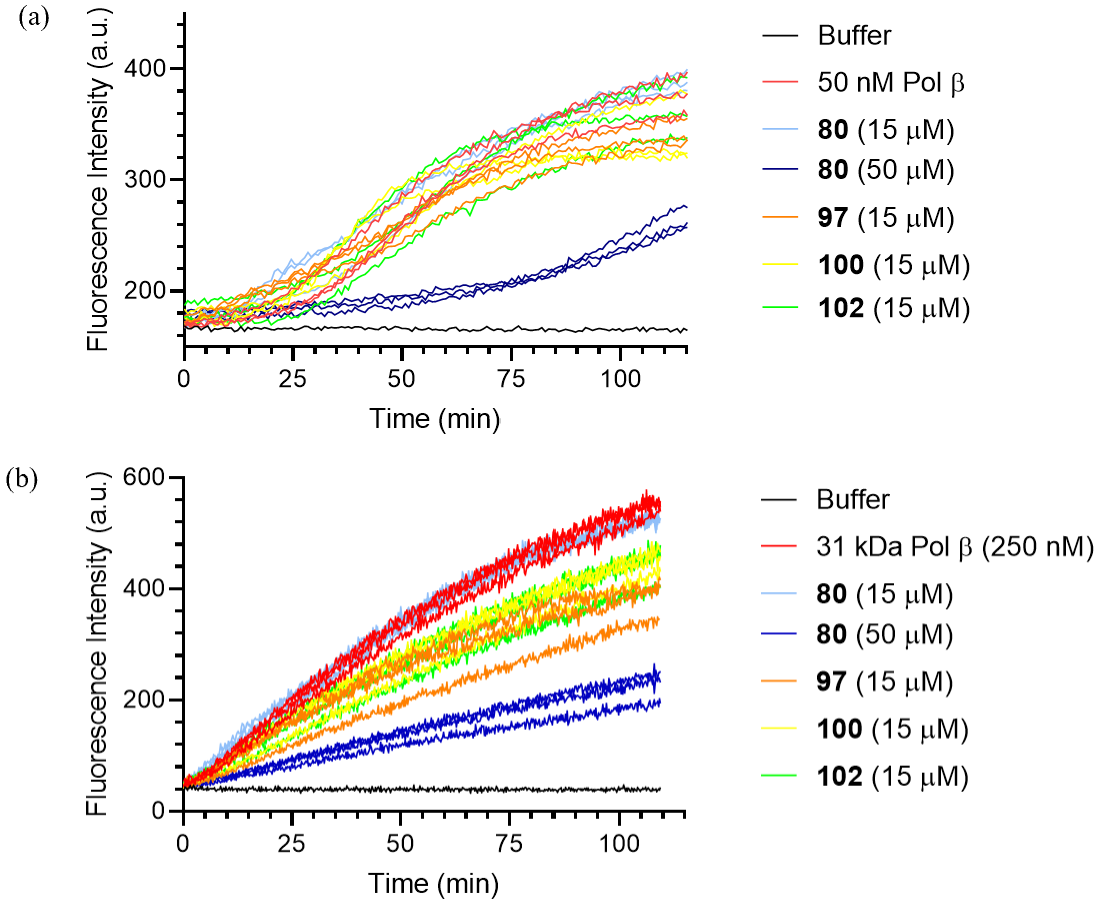


Figure 74. Inhibition of strand displacement activity in the intact enzyme (a) and the 31 kDa domain (b). Pol β (100 nM) or 31 kDa Pol β (250 nM) was preincubated with hit compounds (15 or 50 μM) for 20 min. The preincubation mixture was diluted 10-fold and mixed with DNA cocktail solution, to yield 5 nM Pol β, 50 nM DNA substrate **B,** and 100 μM dTTP**.**

We suspect that we did not identify any potent, selective mutant inhibitor because our second-generation library was based off a compound that binds to the polymerase domain. We think that it is still possible to identify selective inhibitors for a mutant form of Pol β by generating the point mutations in the separate domains using site-directed mutagenesis (Section 4). We envision multiple approaches:

1. Distinct domains containing the cancer-associated point mutations must be generated using site-directed mutagenesis. This will represent a definitive way to confirm mutant

selectivity using kinetic assays (Sections 3.2).

2. Targeting lyase mutants: One could start with a first-generation compound that is a confirmed lyase domain binder (e.g. **47**, Section 3.3.1) and generate a second-generation library with the hydroxylated naphthoic acid already installed on the nucleobase. One could then screen the second-generation library for any molecules that preferentially inhibit a lyase mutant.
3. Targeting polymerase mutants: Several mutants have been identified in the polymerase domain.[13] Therefore, one could take the same approach and synthesize and screen a first-generation library against WT and polymerase mutants. From the hits identified, a second-generation library can be built from the most promising candidates.

Regardless of the domain targeted, using the separate WT and mutant domains to confirm the location of binding before proceeding to a second-generation compound will be a more prudent approach.

3.6 Synthetic lethality induced by Pol β inhibition

As mentioned earlier, DNA repair inhibition is an attractive route at enhancing the cytotoxicity of DNA damaging agents (Section 2.4).[14–19] We were hopeful that our work targeting a mutant form of Pol β would result in a compound that could selectively kill cancer cells, while leaving noncancerous cells alive (Section 3.5). The caveat of targeting WT Pol β is that it is constitutively expressed in all cells.[176–178] By inducing synthetic lethality, one can produce selectivity for cancer cells, despite the target being expressed in all healthy cells. To review, synthetic lethality occurs when the loss or disruption of two genes results in cell death, whereas the loss of only gene in the pair does not (Section 2.4).

Inducing synthetic lethality in cancer cells that lack a particular pathway is a great way to exploit the cancer cell's inherent weakness. An example of this is PARPi treatment in BRCA1-deficient cancer cells.[136,137] HR-deficient cells are more susceptible to PARP inhibitors due to a synthetic lethal interaction. Synthetic lethality is advantageous because it does not require the use of a DNA damaging agent. Rather, it can be induced by a monotherapy in repair-deficient cancer cells without affecting repair-competent noncancer cells. A synthetic lethal interaction between Pol β and BRCA-deficiency has been proposed for several years but has never been confirmed or used clinically due to the lack of therapeutically relevant Pol β inhibitors.[12,69,101]

To evaluate synthetic lethality between Pol β and BRCA1, we treated breast cancer cell lines: MCF-7 (WT, Figures 66-67, Section 3.4.2.3) and MDA-MB-436 (BRCA1-deficient) with **pro-72**. The cytotoxicity of **pro-72** to MDA-MB-436 cells was dependent on concentration and treatment time (Figure 75). In addition, it was evident that **pro-72** was significantly more toxic to BRCA1-deficient cells (Figure 75) than WT MCF-7 cells (Figure 66) and WT HeLa cells (Figure 64). For example, a 1 h treatment of MCF-7 cells with **pro-72** (5 μM) killed about 10% of the cells but the same treatment in MDA-MB-436 resulted in 42% cell death (Figure 75, Table 5). Furthermore, a 2 h treatment with **pro-72** (5 μM) killed 70% MDA-MB-436 cells, whereas the same treatment killed slightly more than half of that (i.e. 39%) in MCF-7 cells (Figures 66 and 75, Table 5). Altogether, this was a significant finding because it supported the proposed synthetic lethal interaction that occurs due to the loss of Pol β function and the BRCA1 gene.

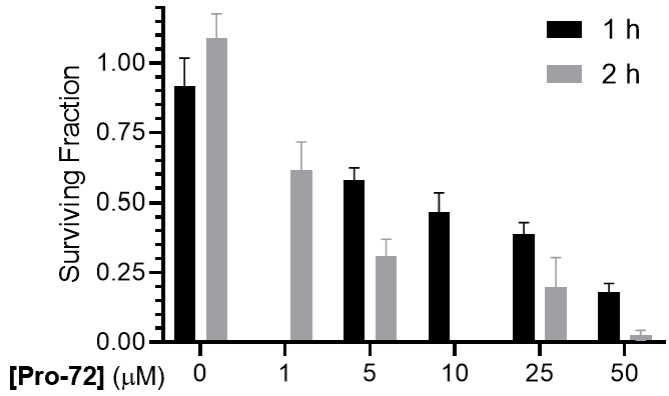


Figure 75. Clonogenic assay of **pro-72** toxicity in MDA-MB-436. MDA-MB-436 cells were treated with **pro-72** (various concentrations) for 1 or 2 h. After treatment, media containing **pro-72** was removed, cells were harvested, counted, and seeded to allow for colony growth. Colonies were visualized by crystal violet.

Table 5. Cell death induced by **pro-72** in MCF-7 and MDA-MB-436 cells.

Pro-72 Treatment	MCF-7	MDA-MB-436
1 μM, 2 h	9%	40%
5 μM, 1 h	10%	42%
5 μM, 2 h	39%	70%

In addition to the higher toxicity of **pro-72** in BRCA1-deficient cells, a slight synergistic effect between our pro-inhibitor and MMS was observed. As these cells are generally less healthy than WT cells, we used gentler conditions (i.e. 0.1 mM MMS). We observed a 1.4-fold enhancement of cell death for combination treatment of MMS (0.1 mM) and **pro-72** (5 μM) for 1 h (Figure 76). Under harsher conditions (i.e. longer treatments with 0.1 mM MMS, 0.2 mM MMS), cell death was too extensive to observe a synergistic effect. Unsurprisingly, treatment with BLM was extremely toxic to BRCA-deficient cells (Figure 76). BLM induces DSBs and MDA-MB-436 cells are HR-deficient and, therefore, less successful at repairing damage induced by BLM.

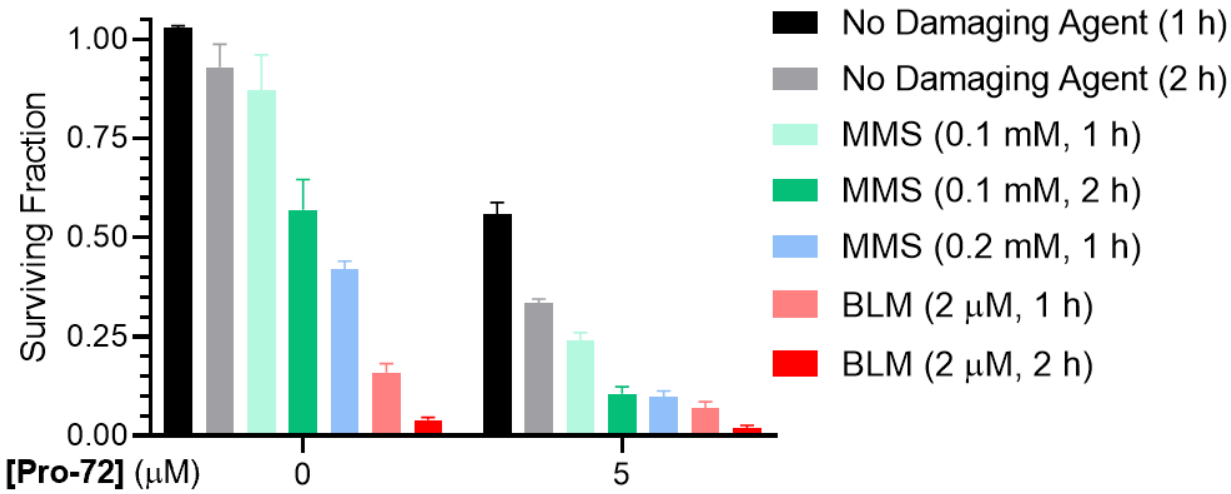


Figure 76. Synergistic effect of **pro-72** with DNA damaging agents on cell death in MDA-MB-436. MDA-MB-436 cells were treated with MMS (0.1, 0.2 mM) or BLM (2 μM) in the presence or absence of **pro-72** (various concentrations) for 1 or 2 h. After treatment, media containing **pro-72** was removed, cells were harvested, counted, and seeded to allow for colony growth. Colonies were visualized by crystal violet.

To confirm the synthetic lethal interaction we observed between **pro-72** and BRCA1-deficiency, we determined:

1. The effect of knocking down Pol β in BRCA1-deficient cells (MDA-MB-436).
2. How **pro-72** compares to a PARPi (Olaparib) in BRCA1-deficient cells (MDA-MB-436).

To address the first issue, we turned to RNA interference to knock down Pol β in MDA-MB-436 cells. We attempted to do this using Pol β shRNA because shRNA is packaged in a DNA plasmid which has several advantages, including being easy to handle and amplifiable in bacterial cells. However, transfection into cells using the suggested transfection reagent (FuGene 6) resulted in extensive (> 95%) cell death, attributed to poor transfection efficiency.[179] The shRNA plasmid possessed resistance to puromycin as a selectable marker. After shRNA transfection, puromycin was added to media to yield only cells that were transfected by the plasmid. When puromycin was not used as a selectable marker, the cells harvested after shRNA transfection retained WT levels of Pol β, determined by western blot analysis.

Therefore, we turned to siRNA to knock down Pol β. Our first transfection with Pol β siRNA using Lipofectamine RNAiMAX was partially successful, resulting in about 50% knock down in Pol β. After increasing the ratio of siRNA:cells to 100 pmol:10^6 cells, we achieved a 70-75% KD in Pol β expression (Figure 77), which was comparable to what is observed in the literature.[179] The use of an internal standard (i.e. H3) allowed us to quantify the efficiency of Pol β knock down in cells.

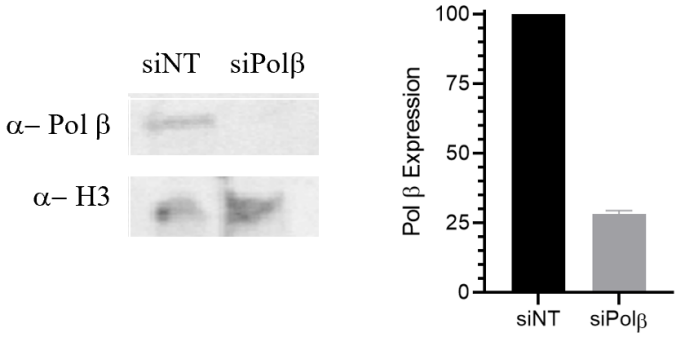


Figure 77. Pol β KD in MDA-MB-436 cells. siNT = RISC-free non-targeting control siRNA, siPolβ = siRNA targeting PolB gene. MDA-MB-436 cells were treated with either control siRNA (siNT) or Pol β siRNA (siPolβ) using two consecutive transfections of Lipofectamine RNAiMAX and siRNA (100 pmol/10^6 cells). Cells were transfected for 72 h, harvested, and reseeded for a clonogenic assay. A portion of transfected cells (1-2 x 10^6) were separate for Western Blot analysis.

Cell viability assays using BRCA1-deficient control (solid) and Pol β KD (dashed) cells confirm several things (Figure 78). First, similar toxicity (~50%) was observed in the cells treated with a non-targeting siRNA control (siNT) in the presence of **pro-72** (5 μM) as previous biological replicates (Figures 75-76), confirming these experiments were consistent and reproducible. In addition, the treatment of **pro-72** at higher concentration (25 μM) in control cells resulted in more cell death. Secondly, knocking down Pol β in MDA-MB-436 cells was toxic and killed roughly 50% of cells. Knocking down Pol β in MDA-MB-436 cells yielded similar cell death to MDA-

MB-436 cells treated with **pro-72** (i.e. 50% cell death). This supported our hypothesis that the loss of Pol β (either by inhibition or KD) was synthetic lethal in BRCA1-deficient cells. In contrast to the control cells, Pol β KD cells exhibited no additional cell death in response to treatment with **pro-72**. This suggested that **pro-72** was not toxic in cells lacking Pol β.

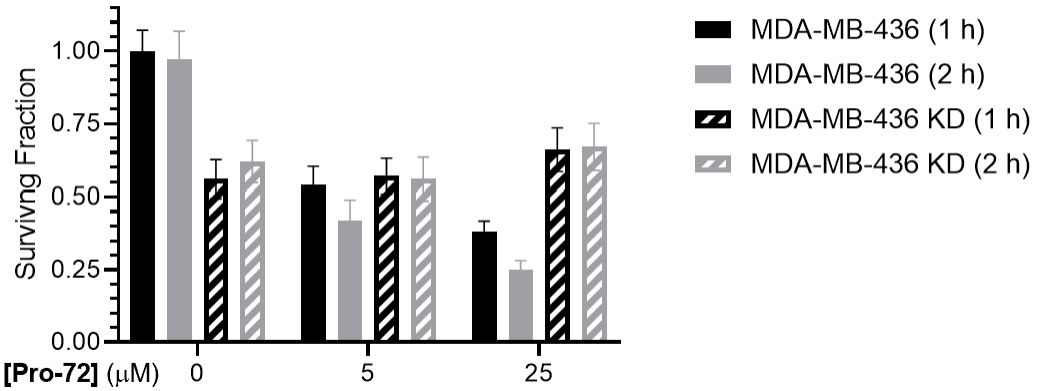


Figure 78. Effect of knocking down Pol β in MDA-MB-436 cells. MDA-MB-436 control and Pol β knock down cells were treated with **pro-72** (various concentrations) for 1 or 2 h. After treatment, media containing **pro-72** was removed, cells were harvested, counted, and seeded to allow for colony growth. Colonies were visualized by crystal violet.

Our second objective in this study was to compare the synthetic lethality of Pol β in BRCA1-deficient cells with the well-known synthetic lethality of PARPi in BRCA1-deficient cells (Section 2.4). Olaparib is a well characterized PARPi with FDA approval for the treatment of BRCA-mutated ovarian, breast, pancreatic, and prostate cancer (Figure 19).[136,137] Olaparib is typically used in cells at 0.1-1 μM.[169] Therefore, MCF-7 and MDA-MB-436 cells were treated with Olaparib at 0.5 and 1 μM for 1 and 2 h (Figure 79). We observed slight toxicity in MCF-7 cells in response to Olaparib at either concentration or treatment time, resulting in less than 30% cell death. In MDA-MB-436 cells, greater cell death was observed (0.5 μM, ~65% cell death; 1 μM, ~75-80% cell death). We did not observe a significant increase in toxicity in response to longer treatment times at either concentration of Olaparib.

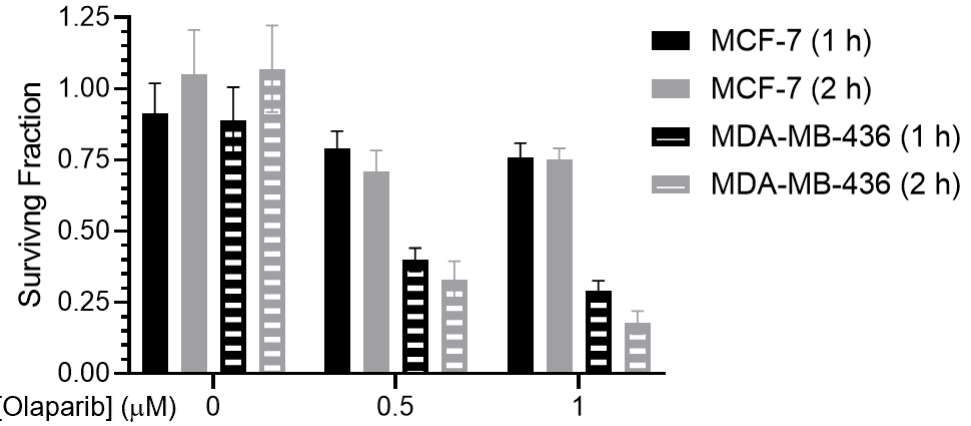


Figure 79. Effect of Olaparib in MCF-7 and MDA-MB-436 cells. MCF-7 and MDA-MB-436 cells were treated with **pro-72** (various concentrations) for 1 or 2 h. After treatment, media containing **pro-72** was removed, cells were harvested, counted, and seeded to allow for colony growth. Colonies were visualized by crystal violet.

In comparison to **pro-72**, Olaparib was a more effective inhibitor at inducing synthetic lethality. For example, to achieve 65-70% cell death in MDA-MB-436, a 2 h treatment of **pro-72** (5 μM) was necessary, a 10-fold decrease in potency compared to Olaparib (Figures 75 and 79, Table 6). On another note, a 2 h, 1 μM treatment of Olaparib resulted in 80% cell death, whereas a 2 h, 1 μM treatment of **pro-72** yielded 40% cell death (Table 6). When comparing the treatments (2 h, 1 μM inhibitor), the enhancement observed in BRCA1-deficient cells by **pro-72** (4.4-fold) was greater than for Olaparib (3.3 fold)! In addition, across the treatments, **pro-72** was less toxic than Olaparib in MCF-7 cells. It is worth mentioning that we did not expect **pro-72** to be better at inducing synthetic lethality than Olaparib. However, this experiment is a nice comparison and shows that a small molecule inhibitor of Pol β has the potential to induce similar effects in a BRCA1-deficient cell line.

Table 6. Percentage of cell death induced by Olaparib and **pro-72** in MCF-7 and MDA-MB-436 cells.

Treatment (2 h)	MCF-7	MDA-MB-436
Olaparib (0.5 μM)	29%	67%
Olaparib (1 μM)	25%	82%
Pro-72 (1 μM)	9%	40%
Pro-72 (5 μM)	39%	70%

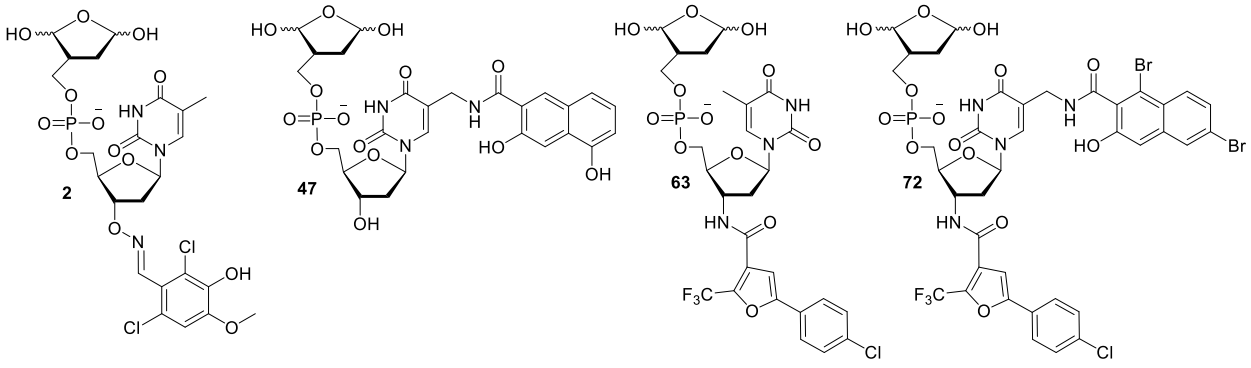
The Pol β pro-inhibitor successfully induced synthetic lethality in BRCA1-deficient cells. We confirmed that knocking down Pol β by siRNA produced similar results in MDA-MB-436 cells (Figure 78). This result is very thrilling because synthetic lethality makes Pol β a better therapeutic target and more selective for cancerous cells.

To complete this project, we will re-evaluate for a synergistic cell killing effect of **pro-72** with damaging agent(s) in MCF-7 and MDA-MB-436 cells using gentler conditions (e.g. lower concentrations of DNA damaging agent(s), shorter treatment time, etc.). We suspect that using too harsh conditions resulted in too much toxicity to determine an accurate enhancement of cell death.

In addition, it has been shown that the loss of Pol β function enhances cell sensitivity to PARPi (Section 2.4).[15] Therefore, it would be interesting to determine the amount of cell death in MCF-7 in response to combination treatment of the two DNA repair inhibitors (i.e. **pro-72** and Olaparib).

4. Future Considerations

We discovered small molecules that are selective for distinct domains of Pol β (Section 3.3.3). In some cases, we have identified small molecules (i.e. **63**, **72**) that selectively bind and inactivate the polymerase domain. In other cases, we identified compounds (i.e. **2**, **47**) that selectively bind to either the 31 kDa or the 8 kDa domain but regardless of the binding location, inactivate lyase and polymerase activity. While we know that our inhibitors interfere with DNA binding, we have been unable to elucidate the underlying mechanisms of how these compounds inactivate a single or both activities. We hypothesize there is a conformational change specific to certain binders that destabilize the enzyme and prevent DNA binding in both domains, thereby inhibiting both lyase and polymerase activity, but that this conformational change is only present in the holo-enzyme.



Inhibitor **47** has been sent to a collaborator for x-ray analysis. It will be interesting to potentially observe the lysine adduct(s) formed from covalent modification by **47** and any corresponding conformational change in the Pol β structure as a result of the inhibitor. We propose using solution NMR to study protein dynamics of Pol β in response to **72** (selective polymerase binder and inhibitor) and **2** (selective polymerase binder but inactivates both activities). We

anticipate a conformational change occurs upon binding to **2** that does not occur upon binding to **72**. This would explain the double inactivation of inhibitor **2** and not in **72**.

The second-generation Pol β WT inhibitor (**72**) has been extensively studied biochemically and in cells. We established its *in vitro* selectivity for Pol β over other polymerases. We plan to use WT MEFs as a comparison to HeLa cells to understand if our inhibitor has any selectivity between healthy and cancerous cells. In addition, WT MEFs and Pol β-null MEFs will be used to establish intracellular selectivity for Pol β.

In addition to Pol β-null MEFs, knock down of Pol β by siRNA in MCF-7 cells can determine selectivity *in vivo*. Using the same method used to knock down Pol β in MDA-MB-436 cells, we knocked down Pol β to 30% relative expression in MCF-7 cells. We expect that if **pro-72** is selective for Pol β, the inhibitor will be less toxic and less effective at killing cells synergistically in cells lacking Pol β. Treatment of MCF-7 Pol β KD cells for 1 h with **pro-72** does not exhibit nonselective toxicity, even at high concentrations (50-100 μM) (Figure 80). Furthermore, at lower concentrations that are typically used for cell viability assays (5 μM), there is no difference in toxicity between control cells and Pol β KD cells. Therefore, MMS sensitivity must be characterized in Pol β KD cells before determining synergism (or the lack of synergism). We anticipate there will be no synergistic effect observed between our pro-inhibitor (**pro-72**) and MMS in the KD cells.

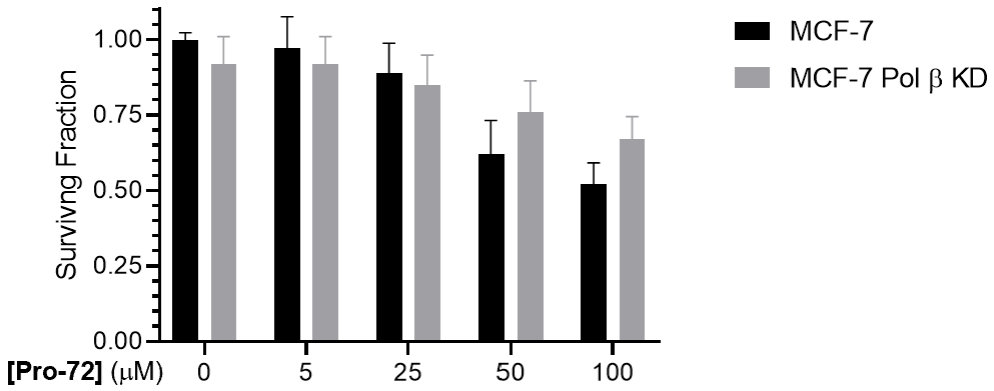
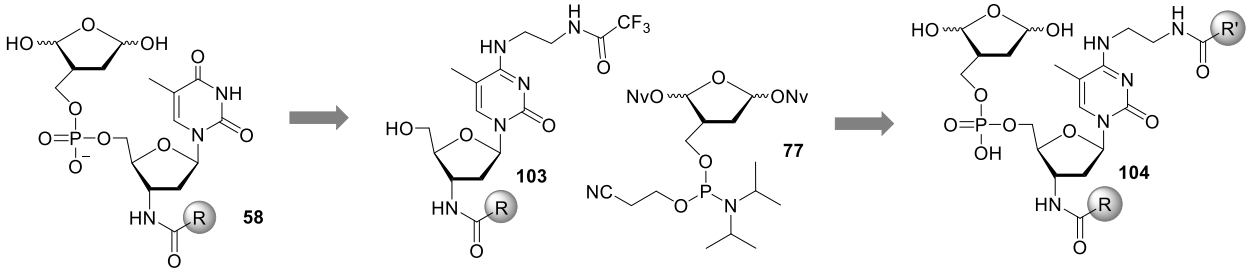


Figure 80. Toxicity of **pro-72** in MCF-7 Pol β KD cells. MCF-7 control and Pol β knock down cells were treated with **pro-72** (various concentrations) for 1 or 2 h. After treatment, media containing **pro-72** was removed, cells were harvested, counted, and seeded to allow for colony growth. Colonies were visualized by crystal violet.

There is great potential for selective inhibitors that target mutant forms of Pol β. As mentioned earlier, we expect this can be successful via two approaches: lyase and polymerase mutants (Section 3.5.3). The most important aspect of identifying selective inhibitors for a mutant will be to generate the point mutation in the distinct domain using site-directed mutagenesis. For example, generating the N24K mutant of the 8 kDa domain can confirm any inhibitors determined when screening the intact N24K Pol β mutant. The same can be done for any polymerase domain mutants screened.

Alternatively, to identify inhibitors for lyase mutants, we propose a new synthetic route for the diversification on the nucleoside in a second-generation molecule. Huijin Lee has designed an improved method for synthesizing libraries containing structural diversity at the C4-position of a pyrimidine ring (Scheme 37). By using a dC template and modifying the exocyclic amine (**103**), both the yield and number of steps in the trifluoroacetamide synthesis at this position could be improved (Schemes 19 and 21).

Scheme 37. Proposed second-generation mutant analog of **47**.



After a first-generation inhibitor (R, Scheme 37) is identified from library **58**, we expect that this synthesis would be more straightforward and result in higher yields than previously used routes that modify the 5-methyl position of dT. Of course, it is likely that the different orientation will affect interactions between the inhibitors and Pol β. We already established that changing the oxime to the amide yields very different inhibitor hits (Section 3.3.3).

Upon synthesis of the new scaffold, the library compounds can be coupled at the C4-ethylene diamine (**104**) and screened against the lyase mutants (N24K, P63L, and T79I) and Pol β WT using the UBER assay to detect any hit compounds that selectively inactivate a mutant. In the proposed second-generation library, we hope that the bottom portion (R) of the molecule could direct the library to the lyase domain, and the top portion (R') could establish unique binding interactions with mutants to improve selectivity and potency.

Most of the inhibitors we have identified (i.e. **2, 63, 72**) target the polymerase domain. Therefore, we propose the same approach of synthesizing the first-generation library (**58**) and screening for polymerase domain mutants. Polymerase domain variants (40) were isolated from colon cancer tumors.[13] Mutants in the DNA binding region that retain normal function are of particular interest. These include R112W, H134Y, H135Y, Q136X*, and T233A* (*Normal or reduced activity was not determined) (Figure 81).[13] A second-generation scaffold can be built from

the first-generation hit using the same method to identify mutants in the polymerase domain (Scheme 37).

Regardless of the approach, we suggest utilizing the separate WT and generated mutant domains to evaluate selectivity of any hit compounds identified. We are hopeful that this will confirm the inhibitor is binding the desired mutant domain prior to exhaustively characterizing the inhibitor *in vitro*.

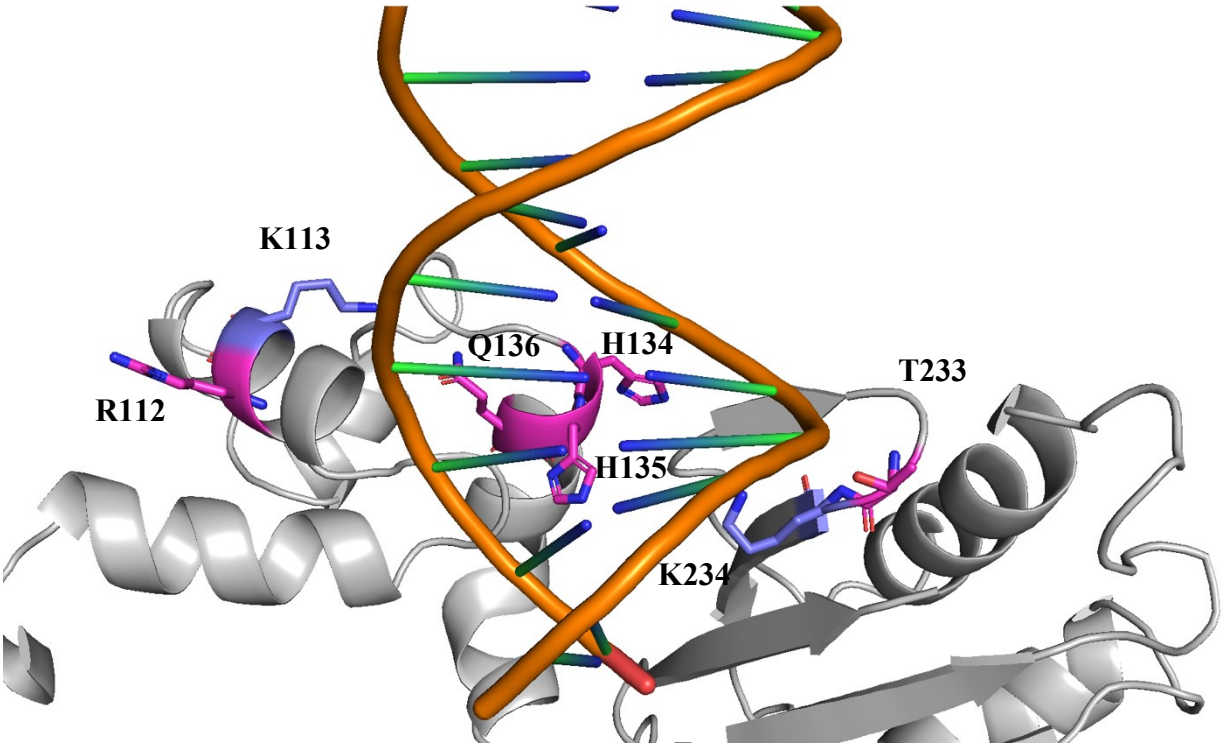


Figure 81. Amino acid (pink) positions of colon cancer-associated Pol β mutants and their proximity to lysines modified (purple) by covalent inhibitor (**72**) (PDB: 1BPX).

5. Experimental Methods

General Methods

Modified oligonucleotides were synthesized on an Applied Biosystems Incorporated 394 oligonucleotide synthesizer. Oligonucleotide synthesis reagents including 5'-phosphorylation reagent (Solid CPR II), SIMA HEX (dichloro diphenyl fluorescein) phosphoramidite, THF abasic site analogue (dSpacer), TAMRA phosphoramidite, and BHQ phosphoramidite were purchased from Glen Research (Sterling, VA). Oligonucleotides containing only native nucleotides were purchased from Integrated DNA Technologies (Coralville, IA) and were purified by 20% denaturing polyacrylamide gel electrophoresis (PAGE) (See below). Oligonucleotides were characterized using a Bruker AutoFlex III Maldi-TOF/TOF system.

All chemicals were purchased from Sigma Aldrich, Fisher, or Alfa and were used without further purification. Small quantities of all library compounds (Appendix 1) were purchased from Sigma Aldrich but were from a variety of vendors (e.g. Enamine, CombiBlocks, ChemApex, etc.). Pol η was purchased from EnzyMax. Sybr Gold was purchased from ThermoFisher. Trypsin, dNTPs, terminal deoxynucleotide transferase, Klenow exo$^-$, and T4 polynucleotide kinase were obtained from New England Biolabs. Radionuclides were from Perkin Elmer. Poly-Prep columns were from BioRad. C18-Sep-Pak cartridges were obtained from Waters. Zip-Tips were purchased from Millipore. Photolyses were carried out in a Rayonet photoreactor fitted with 16 lamps having a maximum output at 350 nm. Quantification of radiolabeled oligonucleotides was carried out using a Molecular Dynamics Phosphorimager 840 equipped with ImageQuant.

Tryptone and yeast extract were from BD Biosciences. Protease inhibitor cocktail (EDTA-free) was from Roche. HisTrap column, Heparin column, HiTrap Q column, HiTrap SP column, and Q Sepharose fast flow were from GE Healthcare. Affi-gel blue gel resin was from BioRad. BL21 (DE3) pLysS and Rosetta 2 DE3 *E.coli* were from Millipore. BL21 (CodonPlus) DE3-RIL and DH5α *E.coli* were from ThermoFisher. Maxi Prep and Mini Prep Kits were from Qiagen. Sonication was conducted using a Brandson SFX-150 sonifier. Protein purification was conducted using an AKTA FPLC. All centrifugations during protein preparation were done using a Beckman Coulter Allegra25r.

UPLC-MS/MS analyses were carried out on a Waters Acquity/Xevo-G2 UPLC-MS system equipped with an ACQUITY UPLC HSS T3 Column (100 Å, 1.8 μm, 2.1 mm x 100 mm). Masses were obtained via deconvolution using MassLynx 4.2 software or BioPharmaLynx 1.3.2 software.

Well plates used for organic solvents and photolyses were obtained from VWR. Well plates used for fluorescence assays were obtained from Corning (CLS3825, 384 well plates, for homogenous luminescent and HTRF assays). Fluorescence data were collected on a Varian Cary Eclipse fluorescence spectrophotometer equipped with a well plate attachment or a CLARIOstar Plus microplate reader at the Center for Molecular Biophysics at Johns Hopkins University. Fluorescence anisotropy measurements were conducted using an AVIV Biomedical Model ATF 107 spectrofluorometer at the Center for Molecular Biophysics at Johns Hopkins University.

The following items were generous gifts from colleagues. Mutant plasmids (Pol β N24K, Pol β P63L, and Pol β T79I) were received from Professor Joann Sweasy, University of Arizona Cancer Center. Plasmids for the 8 kDa and 31 kDa domains were from Dr. Sam Wilson, NIH. Pol λ plasmid was from Professor Zucai Suo, Florida State University. The UBER probe precursor (CCVJ) was a gift from Professor Eric Kool, Stanford University. MCF-7 and MDA-MB-436 cell lines were acquired from Professor Theodore DeWeese, Johns Hopkins University.[160]

Dulbecco's Modified Eagle Medium (DMEM) with high glucose, RPMI 1640 Medium with GlutaMAX supplement, Opti-MEM reduced serum medium, and insulin were obtained from ThermoFisher. Antibiotic antimycotic solution (penicillin, streptomycin, and amphotericin B), and fetal bovine serum (FBS) were obtained from MilliporeSigma. PBS buffer was obtained from Quality Biological. Pre-cast 4-20% SDS-PAGE, Precision Plus Protein WesternC standard, Western blot kit, and StrepTactin-HRP conjugate were from BioRad. NP-40 was obtained from Sigma Aldrich. All siRNAs (siGENOME RISC-Free Control, D-001220-01; siGENOME Human POLB siRNA, D-005164-04) were purchased from Horizon Discoveries. All antibodies (Recombinant Anti-DNA Polymerase beta antibody, ab175197) were purchased from Abcam. Cells were counted using a BioRad TC20 cell counter. Western blots were carried out using a BioRad Trans-Blot Turbo Transfer system. Ponceau red stain was prepared with ponceau S tetrasodium salt (0.1%, Sigma Aldrich) and acetic acid (5%) in distilled H_2O. Developing reagents were acquired from Thermo. Western blots were visualized using a Typhoon 9410 equipped with chemiluminescence imaging at the Integrated Imaging Center at Johns Hopkins University.

All small molecules synthesized were characterized using a Bruker Avance 400 MHz Spectrometer or a Varian Inova 800 MHz spectrometer at Johns Hopkins University.

Synthesis, Purification, and Characterization of Modified Oligonucleotides

An oligonucleotide containing a 5'-dRP photolabile precursor was synthesized using standard phosphoramidites and protocols from Glen Research.[29] The coupling step involving the AP precursor lasted 5 min and the phosphorylating agent (Solid CPR II) was double-coupled for 5 min each cycle.

Oligonucleotides were deprotected using a solution (0.8 mL) of concentrated ammonium hydroxide (400 μL) and 40% methylamine in H_2O (400 μL) at 65 ºC for 1 h. Following deprotection, oligonucleotides were dried in a Speed Vac concentrator, resuspended in loading buffer (95% formamide with 10 mM EDTA), purified by 20% denaturing polyacrylamide gel electrophoresis (PAGE), and visualized using a TLC plate and UV lamp. Oligonucleotides containing photochemical precursors were exposed to UV light for the shortest time possible and were always covered with foil, whether the DNA was on the gel or in a tube. The band was cut from the gel, crushed into small pieces, and eluted from the gel matrix with elution buffer (2-10 mL, 100 mM NaCl, 1 mM EDTA) for 12 h with shaking. The solution was filtered using a Poly-Prep column and desalted using C18-Sep-Pak cartridges.

All modified oligonucleotides in this thesis were previously characterized by MALDI-TOF-MS (Table 1, Section 3.2.4).[26,27,135] For MALDI-TOF-MS, oligonucleotides (100 pmol) were desalted by Zip-Tip and resuspended in matrix (1 μL) containing 3-hydroxypicolinic acid (10 mg/mL) and ammonium citrate (30 mg/mL).

Radiolabeling and Preparation of Oligonucleotide Complexes

Oligonucleotides were 3'-^{32}P labeled by α^{32}–P Cordycepin triphosphate and terminal deoxynucleotidyl transferase. In addition, 3'-^{32}P labeling was also completed using Klenow exo$^-$ and α-^{32}P-ATP, within a duplex containing a single 5'-dT overhang. In this case, the 3'-terminus of the 5'-dRP strand was ^{32}P-labeled, denatured from the duplex, and purified by 20% denaturing PAGE. Labeling with 5'-^{32}P labeling was completed by T4 polynucleotide kinase and γ-^{32}P-ATP. Ternary complexes were hybridized by mixing ^{32}P-labeled oligonucleotides with the appropriate template and flanking strand in a 1:2.5:5 ratio in phosphate buffered saline (10 mM sodium phosphate, 100 mM NaCl, pH 7.3), heating to 95 ºC, and slowly cooling to 25 ºC.

Ternary complexes containing fluorophore labeled oligonucleotides were prepared by annealing the fluorophore-labeled strand with the appropriate quencher-labeled template and flanking strand in a 1:2:3 ratio.

Bradford Assay

Bradford assays were used to determine the concentration of purified proteins and proteins in cell lysates. A standard curve was generated by adding Bradford reagent (950 μL, Sigma Aldrich) to solutions (50 μL) containing 50, 100, 250, 500, 750, 1000 mg/mL bovine serum albumin (BSA). After incubating for 10 min at 25 ºC, the absorbance (595 nm) of each standard was measured using a spectrophotometer. Absorbance was plotted against amount of BSA and fit to a straight line. An aliquot of purified protein (5 μL) was diluted 10-fold (50 μL) and mixed with Bradford reagent (950 μL). After incubating for 10 min, the absorbance (595 nm) was used to determine the concentration of protein based on the standard curve equation.

Amplification of plasmid containing the 8 kDa fragment of Pol β[180]

Luria broth (LB) Media and LB Agar were prepared and autoclaved to make plates. LB Media was prepared by dissolving 5 g Bacto-tryptone, 2.5 g Bacto-yeast, and 5 g NaCl in 500 mL H_2O. LB Agar was prepared by dissolving 5 g Bacto-tryptone, 2.5 g Bacto-yeast, 7.5 g Agar and 5 g NaCl in 500 mL H_2O. These solutions were autoclaved for 25 min. While the LB Agar was still warm, an aliquot of ampicillin stock solution (1000X, 100 mg/mL ampicillin, 50 μL) was sterilized using a sterile filter and added to the 500 mL LB Agar. In a sterile environment, the mixture was poured into plates and allowed to cool.

DH5α competent cells (40 μL) were thawed at 0 ºC. pRSET plasmid containing the coding region for the 8 kDa Domain of Pol β (80 ng/μL, 1 μL) was added to the tube containing DH5α cells and incubated at 0 ºC for 30 min. After 30 min, the cells were placed in a 42 ºC water bath

for 45 s and immediately put back on ice for 2 min. SOC outgrowth medium (200 μL, 5x) was added to the cells. The medium shook at 37 ºC for 1 h to allow the cells to recover. After 1 h of recovery time, multiple diluted cells samples (1:10, 1:50, and 1:100) were plated onto Agar plates containing ampicillin.

The plates incubated at 37 ºC overnight. After colonies had grown (~14 h), the plates were stored at 4 ºC. An overnight culture was prepared by mixing an aliquot of LB media (10 mL) and ampicillin stock solution (1000X, 100 mg/mL, 10 μL). One colony was removed from the DH5α (1:100 dilution) plate and placed in the medium. The cells were resuspended and shaken at 37 ºC overnight. After 14 h, the cells were spun down at (4 ºC, 10 min, 4500 rcf).

The Qiagen mini prep was used to purify the plasmid DNA. The steps were as follows: The supernatant was removed and replaced with 500 μL of Resuspension P1 buffer with RNase A. The cells were resuspended and divided into two 250 μL aliquots. To each mini prep sample, P2 Lysis Buffer, pH 11 (250 μL) was added to break down the cell walls and the tube was inverted several times. To prevent DNA damage, P3 Neutralization buffer containing NaOAc, pH 4 (300 μL) was added until the mixture was homogeneous. The tubes were spun down for 10 min at 4 ºC. The supernatant was collected and passed through the blue mini prep column. The flow-through was thrown away. PE Buffer with EtOH (750 μL) was added to the column, incubated for 2 min, and spun down for 1 min at 4 ºC. The flow-through was removed and this was repeated once more with PE Buffer. After the second wash with PE Buffer, the column was spun down twice for 2 min each. Amplified plasmid was eluted with 10 mM Tris buffer EB (50 μL) and was spun down for 1.5 min at 4 ºC. The flow-through was collected and stored. A small portion (2 μL) was used to determine the concentration (800 ng/μL) using a NanoDrop ($\varepsilon_{260} = 1.1\times10^7\ M^{-1}cm^{-1}$).[180]

Transformation and Expression of 8 kDa Pol β domain[180,181]

LB media and LB Agar were prepared and autoclaved to make plates. LB Media was prepared by dissolving 5 g Bacto-tryptone, 2.5 g Bacto-yeast, and 5 g NaCl in 500 mL H_2O. LB Agar was prepared by dissolving 2.5 g Bacto-tryptone, 1.25 g Bacto-yeast, 3.75 g Agar and 2.5 g NaCl in 250 mL H_2O. These solutions were autoclaved for 25 min. While the LB Agar was still warm, an aliquot of ampicillin stock solution (1000X, 100 mg/mL ampicillin, 25 μL) and an aliquot of chloramphenicol (1000X, 34 mg/mL chloramphenicol in EtOH, 25 μL) were sterilized with a sterile filter and added to the 250 mL LB Agar. In a sterile environment, the mixture was poured into plates and allowed to cool.

BL21 DE3 pLysS competent cells (20 μL) were thawed at 0 °C. pRSET plasmid containing the coding region for the 8 kDa domain of Pol β (80 ng/μL, 1 μL) was added to the tube containing *E.Coli* cells and incubated at 0 °C for 30 min. After 30 min, the cells were placed in a 42 °C water bath for 45 s, and immediately put back in ice for 2 min. SOC outgrowth medium (100 μL, 5x) was added to the cells. The cells were incubated at 37 °C for 1 h with shaking (180 RPM) to allow the cells to recover. After 1 h of recovery time, multiple diluted cell samples were prepared by mixing cells with SOC medium:

1. Not diluted (50 μL cells)
2. 1:10 (5 μL cells + 45 μL SOC medium)
3. 1:50 (1 μL + 50 μL SOC medium)

Each solution was plated onto separate agar plates containing ampicillin and chloramphenicol. The plates incubated at 37 ºC overnight. After colonies had grown, the plates were stored at 4 ºC. One colony was removed from the BL21 DE3 pLysS (1:50 dilution) plate and placed in LB media (10 mL) with ampicillin (1000X, 100 mg/mL, 10 μL) and chloramphenicol (1000X, 34 mg/mL, 10 μL). The cells were resuspended and shaken at 37 ºC overnight.

After 11.5 h, the starter culture was retrieved from the shaker and inoculated by spinning cells down (15,000 RCF, 4 ºC, 10 min), removing the supernatant, and resuspending the pellet in fresh LB media (10 mL). In a sterile environment, culture media was prepared by mixing LB media (2 L), Ampicillin (2 mL, 100 mg/mL), Chloramphenicol (2 mL, 34 mg/mL), and starter culture (10 mL). Cultures were shaken at 37 ºC at 250 RPM until the cells reached an OD of 0.5 (about 4.5 h) At that point, IPTG (1 mL, 1000X 1 M) was added to each batch of cells to make a final concentration of 1 mM IPTG. The cells were shaken for another 3 h at 37 ºC.

After 3 h, the cells were harvested by spinning them down (15,000 RCF, 4 ºC, 10 min) and removing the media. The dry pelleted cells were stored at -80 ºC until purification.

Purification of the 8 kDa domain of Pol β[181]

Buffer A: 50 mM Tris HCl, 1 mM EDTA, pH 7.5, 1 mM PMSF, 0.1 mM protease inhibitor.

Buffer B: 50 mM Tris HCl, 1 mM EDTA, 1 M NaCl, pH 7.5, 1 mM PMSF, 0.1 mM protease inhibitor.

All buffers were supplemented with PMSF (1 mM) and protease inhibitor cocktail (0.1 mM). Cells were thawed and resuspended in lysis buffer (20 mL, 50 mM Tris•HCl, 1 mM EDTA, 500 mM NaCl, 0.1 mM inhibitor cocktail, pH 7.5). The cells were sonicated at 0 ºC for 10 min (10 min total time, 20 s on, 40 s off, 50% power). The cell lysate was centrifuged for 20 min at 15,000 RCF at 4 ºC. Q-sepharose resin (15 mL, 20% EtOH) was spun down and washed with H_2O (2 x

15 mL) and Buffer B (2 x 15 mL), each time discarding the supernatant. The supernatant from the lysis was filtered with a sterile syringe into equilibrated Q-sepharose resin, diluted with Buffer A until 100 mM NaCl, and was shaken at 4 ºC for 2 h. The ssDNA-cellulose resin (500 mg) was washed with H_2O (2 x 10 mL) and Buffer B (2 x 10 mL).

After 2 h, the Q-sepharose resin solution was passed through a gravity Poly-Prep column (20 mL, disposable), at 4 ºC. The flow-through was collected and mixed with ssDNA-cellulose resin at 4 ºC for 1 h. The solution was then passed through a prep column and the protein was eluted by increasing salt concentration (NaCl, 100 mM to 1 M). This purification was completed manually. Buffers containing multiple salt concentrations (i.e. 100 mM, 200 mM, 400 mM, 600 mM, 800 mM, and 1 M NaCl, 10 mL each) were added to the column consecutively. Fractions (3 x 3.3 mL) from each salt concentration were collected and analyzed by SDS-PAGE to determine which fractions contained the 8 kDa domain. The protein eluted at 200-400 mM NaCl, based on SDS PAGE analysis.

The fractions containing the 8 kDa domain of Pol β were dialyzed overnight in Buffer B (2 x 500 mL). The fractions were poured into a SuperLoop and injected onto GE HiTrap SP FF column (5 mL) and eluted by increasing salt concentrations (1 mL/min, NaCl, 100 mM to 1 M in 50 min, 2 mL fractions collected). The protein eluted around 400 mM NaCl, based on UV absorbance and SDS PAGE analysis. These fractions were combined and concentrated using a 3 kDa 500 µL Amicon filter until reaching a concentration of 68 µM in 200 µL (determined by Bradford assay), at which point a solution of glycerol in water (50 µL, 50% v:v) was added to the stock. The final solution consisted of 55 µM 8 kDa Pol β, 10% v:v glycerol, and 250 µL. The protein was aliquoted, flash frozen, and stored at -80 ºC until use.

Transformation and Expression of the 31 kDa domain of Pol β[182]

LB Media and LB Agar were prepared and autoclaved to make plates. LB Media was prepared by dissolving 45 g Bacto-tryptone, 22.5 g Bacto-yeast, and 45 g NaCl in 4.5 L H_2O. LB Agar was prepared by dissolving 2.5 g Bacto-tryptone, 1.25 g Bacto-yeast, 3.75 g Agar and 2.5 g NaCl in 250 mL H_2O. These solutions were autoclaved for 25 min. While the LB Agar was still warm, an aliquot of ampicillin stock solution (1000X, 100 mg/mL ampicillin, 250 μL) and an aliquot of chloramphenicol (1000X, 34 mg/mL chloramphenicol in EtOH, 250 μL) were sterilized in a sterile filter and added to the 250 mL of LB Agar. In a sterile environment, the mixture was poured into plates and allowed to cool.

BL21 DE3 pLysS competent cells (20 μL) were thawed at 0 ºC. pRSET plasmid containing the coding region for the 31 kDa Domain of Pol β (80 ng/μL, 1 μL) was added to the tube containing *E.coli* cells and incubated at 0 ºC for 30 min. After 30 min, the cells were placed in a 42 ºC water bath for 45 s, and immediately put back in ice for 2 min. SOC outgrowth medium (100 μL, 5x) was added to the cells. The cells were incubated at 37 ºC for 1 h with shaking (180 RPM) to allow the cells to recover. After 1 h of recovery time, multiple diluted cell samples were prepared:

1. 1:10 (5 μL cells + 45 μL SOC medium)
2. 1:50 (1 μL + 50 μL SOC medium)
3. 1:100 (0.5 μL cells + 45 μL SOC medium)

Each solution was plated onto separate agar plates containing ampicillin and chloramphenicol. The plates were incubated at 37 ºC overnight. After colonies had grown, the plates were stored at 4 ºC. One colony was removed from the BL21 DE3 pLysS (1:50 dilution)

and placed in LB media (10 mL) with ampicillin (1000X, 100 mg/mL, 10 µL) and chloramphenicol (1000X, 34 mg/mL, 10 µL). The starter culture was shaken at 37 ºC overnight.

After 11.5 h, the starter culture was retrieved from the shaker and inoculated by removing the supernatant and resuspending the pellet in fresh LB medium (10 mL). In a sterile environment, culture media was prepared by mixing LB media (2 x 1 L), Ampicillin (2 x 1 mL, 100 mg/mL), Chloramphenicol (2 x 1 mL, 34 mg/mL), and starter culture (2 x 5 mL). Cultures were shaken at 37 ºC at 250 RPM until the cells reached an OD of 0.5 (about 5 h) At that point, IPTG (1 mL, 1000X 1 M) was added to each batch of cells to make a final concentration of 1 mM IPTG. The cells were shaken for another 3 h at 37 ºC. After 3 h, the cells were harvested by spinning them down (15,000 RCF, 4 ºC, 10 min) and removing the media. The dry pelleted cells were frozen and stored at -80 ºC until purification.

Purification of the 31 kDa domain of Pol β[182]

Buffer A: 50 mM Tris HCl, 1 mM EDTA, 500 mM NaCl, pH 7.5, 10 mM $Na_2S_2O_5$, 1 mM PMSF, 0.1 mM protease inhibitor

Buffer B: 10 mM Tris HCl, 1 mM EDTA, 1 mM PMSF, pH 7.5, 0.1 mM protease inhibitor

Buffer C: 10 mM Tris HCl, 1 mM EDTA, 1 mM PMSF, 500 mM NaCl, pH 7.5, 0.1 mM protease inhibitor

All buffers were supplemented with PMSF (1 mM) and protease inhibitor cocktail (0.1 mM). Cells (~3.5 g) were thawed and resuspended in Buffer A (20 mL, 50 mM Tris HCl, 1 mM EDTA, 10 mM $Na_2S_2O_5$, 500 mM NaCl, 0.1 mM inhibitor cocktail, pH 7.5). The cells were sonicated at 0 ºC for 10 min (10 min total time, 10 s on, 20 s off, 50% power). The cell lysate was centrifuged for 20 min at 15,000 RCF at 4 ºC. Cell lysate was filtered through a sterile filter into dialysis (3.5 K cut-off) bag and dialyzed against Buffer B at 4 ºC (2 x 1 L) for 18 h. Q-sepharose

resin (15 mL, 20% EtOH) was spun down (10 min, 10,000 RCF, 4 ºC) and washed with H_2O (2 x 15 mL) and Buffer B (2 x 15 mL), each time discarding the supernatant. Dialyzed protein was filtered with a sterile syringe into equilibrated Q-sepharose resin and was shaken at 4 ºC for 1.5 h. After binding, the Q-sepharose resin solution was passed through a gravity Poly-Prep column (20 mL, disposable), at 4 ºC. The flow-through was collected (2 x 10 mL) and the protein was eluted manually by increasing the concentration of NaCl (combining Buffer B and C: 0, 100, 200, 300, 400, 500 mM NaCl, 10 mL each). Fractions (3 x 3.3 mL) from each salt concentration were collected and analyzed to determine which fractions contained the 31 kDa domain. Analysis by 12% SDS PAGE indicated the protein eluted from 200-400 mM NaCl. Those fractions were pooled and dialyzed against Buffer B at 4 ºC (2 x 1 L) for 18 h.

Affi-gel blue gel resin (10 mL) was spun down (10 min, 10,000 RCF, 4 ◦C) and washed with H_2O (2 x 10 mL) and Buffer B (2 x 10 mL), each time discarding the supernatant. The dialyzed protein was added to the equilibrated Affi-gel blue gel resin and shaken at 4 ºC for 2 h. The solution was passed through a gravity Poly-Prep column (20 mL, disposable) at 4 ºC and the protein was eluted manually at increasing salt concentrations (combining Buffer B and C: 0, 100, 200, 300, 400, 500 mM NaCl, 10 mL each). Fractions (3 x ~3.3 mL) from each salt concentration were collected and analyzed to determine which fractions contained the 31 kDa domain. Analysis by 12% SDS PAGE confirmed the protein eluted at 100-200 mM NaCl. Those fractions were pooled and concentrated with a 3K Amicon filter (pre-equilibrated with H_2O (2 x 2 mL) and Buffer B (2 x 2 mL)). The concentration was measured by Bradford assay to be 16.5 μM (400 μL). A solution of glycerol in water (100 μL, 50% v:v) was added to the stock. The final solution consisted of 13.2 μM 31 kDa domain of Pol β, 10% v:v glycerol, and 500 μL. The 31 kDa domain/fragment of Pol

β was aliquoted, flash frozen, and stored at -80 ºC until use. Activity was confirmed using the strand displacement assay (Section 3.2.1).

Expression of Pol β mutants (N24K, P63L, T79I)[13]

Pol β mutant pET28a construct plasmids (N24K, P63L, and T79I) were transformed into Rosetta 2 (DE3) cells (same procedure as described above). LB cultures supplemented with 50 μg/mL kanamycin and 34 μg/mL chloramphenicol were grown overnight at 37 ºC, diluted 1:100 in LB media, and grown until reaching an A_{600} of ~0.5. Cultures were cooled down in a cold room to 4 ºC while the incubator cooled to 20 ºC. Cultures were not induced until they were below room temperature (e.g. cold to the touch) and the incubator reached 20 ºC. Protein expression was induced with 1 mM IPTG at 20 ºC overnight. After 16 h, cells were harvested by centrifugation (4 x 15,000 RCF, 4 ºC, 10 min), frozen, and stored at -80 ºC.

Purification of Pol β mutants (N24K, P63L, T79I)[13]

Buffer A: 40 mM Tris HCl, 0.5 M NaCl, 5 mM Imidazole, 1 mM PMSF, 0.1 mM protease inhibitor, pH 8.0

Buffer B: 40 mM Tris HCl, 0.5 M NaCl, 0.5 M Imidazole, 1 mM PMSF, 0.1 mM protease inhibitor, pH 8.0

Buffer C: 40 mM Tris HCl, 1 M NaCl, 5 mM Imidazole, 1 mM PMSF, 0.1 mM protease inhibitor, pH 8.0

All buffers were supplemented with PMSF (1 mM) and protease inhibitor cocktail (0.1 mM). Cells (~3 g) were thawed and resuspended in Buffer A (30 mL). The cells were sonicated at 0 ºC for 10 min (10 min total time, 10 s on, 10 s off, 50% power). The cell lysate was centrifuged (20 min, 15,000 RCF, 4 ºC). The supernatant of the cell lyase was filtered at 4 ºC through a 0.2 μm sterile filter and applied to a pre-equilibrated HisTrap HP column (5 mL) via a SuperLoop.

HisTrap HP column was used with a linear imidazole elution gradient (1 mL/min, 100% Buffer A -> 100% Buffer B in 50 min, 2 mL fractions collected). Fractions containing mutant (eluted at ~125-250 mM imidazole, determined by UV and SDS PAGE analysis) were combined and concentrated to 1 mL with a pre-equilibrated 10K Amicon filter (Pre-equilibration: 2 x 10 mL H_2O; 2 x 10 mL Buffer A; spun at 7,500 RCF, 4 ºC) at 4 ºC and then diluted to 10 mL. The resulting solution was sterilized with a sterile filter and applied to a 5 mL SP-FF column with a SuperLoop. SP-FF column was used with a linear NaCl elution gradient (1 mL/min, 100% Buffer A -> 100% Buffer C in 50 min, 2 mL fractions collected). Fractions containing mutant (eluted at ~0.5 – 0.6 M NaCl) were combined and concentrated. FPLC-purified proteins were analyzed by SDS-PAGE and Coomassie Blue staining. Protein concentration(s) (listed below) were determined using the A_{280} Nanodrop and the extinction coefficient for Pol β (ε = 21,200 $M^{-1}cm^{-1}$).[13]

[Pol β N24K] = 4.0 μM, 2.4 mL

[Pol β P63L] = 41.2 μM, 0.8 mL

[Pol β T79I] = 43.8 μM, 0.64 mL

Glycerol (50% v:v in water ; 600 μL Pol β N24K, 200 μL, Pol β P63L, 160 μL, Pol β T79I) was added to each stock and the mutants (concentrations and volumes listed below) were aliquoted, flash frozen, and stored at -80 ºC. Mutant activities were confirmed using the strand displacement assay (Section 3.2.1).

[Pol β N24K] = 3.2 μM, 3 mL, 10% v:v glycerol

[Pol β P63L] = 33 μM, 1 mL, 10% v:v glycerol

[Pol β T79I] = 35 μM, 0.8 mL, 10% v:v glycerol

Expression of Pol λ[183]

Pol λ pET28 construct plasmids were transformed into BL21 (CodonPlus) DE3-RIL cells. LB cultures supplemented with 50 μg/mL kanamycin and 34 μg/mL chloramphenicol were grown overnight at 37 ºC. After 14 h, the starter cultures were spun down (15,000 RCF, 4 ºC, 10 min), replaced with fresh media, diluted 1:100 with LB media, and grown until reaching an A_{600} of ~0.5. Cultures were cooled to 4 ºC in a cold room and the incubator cooled to 22 ºC. Cultures were not induced until the cultures were below room temperature (e.g. cold to the touch) and the incubator reached 22 ºC. Protein expression was induced with 0.4 mM IPTG at 22 ºC for 6 h. After 6 h, cells were harvested by centrifugation (4 x 15,000 RCF, 4 ºC, 10 min), frozen, and stored at -80 ºC.

Purification of Pol λ[183,184]

Buffer A: 20 mM NaH_2PO_4, 0.5 M NaCl, 1 mM DTT, 10% v:v glycerol, 1 mM PMSF, 0.1 mM protease inhibitor, pH 7.8

Buffer B: 20 mM NaH_2PO_4, 0.5 M NaCl, 1 M Imidazole, 1 mM DTT, 10% v:v glycerol, 1 mM PMSF, 0.1 mM protease inhibitor, pH 7.8

Buffer C: 20 mM NaH_2PO_4, 1 M NaCl, 1 mM DTT, 10% v:v glycerol, 1 mM PMSF, 0.1 mM protease inhibitor, pH 7.8

All buffers were supplemented with PMSF (1 mM) and protease inhibitor cocktail (0.1 mM). Cells (~1 g) were thawed and resuspended in Buffer A (10 mL) supplemented with PMSF (1 mM) and protease inhibitor cocktail (0.1 mM). The cells were sonicated at 0 ºC for 10 min (10 min total time, 10 s on, 10 s off, 50% power). The cell lysate was centrifuged (20 min, 15,000 RCF, 4 ºC). The supernatant of the cell lyase was filtered at 4 ºC through a 0.2 μm sterile filter and applied to a pre-equilibrated 5 mL HisTrap HP column (Pre-equilibration: 2 mL/min; Buffer A, 20 mL; Buffer B, 10 mL; Buffer A, 20 mL) via a SuperLoop.

HisTrap HP column was used with a linear imidazole elution gradient (1 mL/min, 100% Buffer A -> 100% Buffer B in 50 min, 2 mL fractions collected). Fractions containing Pol λ (eluted at ~250-500 mM Imidazole, determined by UV and SDS PAGE analysis) were combined and concentrated to 1 mL with a pre-equilibrated 10K Amicon filter (Pre-equilibration: 2 x 10 mL H_2O, 2 x 10 mL Buffer A; spun at 7,500 RCF, 4 ºC) at 4 ºC and then diluted to 10 mL. The resulting solution was applied to a 5 mL Heparin column with a SuperLoop. The Heparin column was used with a linear NaCl elution gradient (1 mL/min, 100% Buffer A -> 100% Buffer C in 50 min, 2 mL fractions collected). Fractions containing Pol λ (eluted at ~750 mM NaCl) were combined and concentrated. FPLC-purified proteins were analyzed by SDS-PAGE and Coomassie Blue staining. Protein concentration (14.5 μM) was determined using a Bradford assay. All buffers contained 10% v:v glycerol so the final solution consisted of 14.5 μM Pol λ, 10% v:v glycerol, and 400 μL. The protein was aliquoted, flash frozen, and stored at -80 ºC.

Fluorescence-based strand displacement assay (used for Pol β, 31 kDa Pol β, and Klenow exo$^-$)

A 10X preincubation mixture was prepared by mixing a 40X working solution of Pol β (25 μL, 400 nM) with an aliquot of the inhibitor (5 μL, 1 mM in 1:1 MeCN/H_2O, 40X) in 1X reaction buffer (total volume: 100 μL; 50 mM HEPES buffer pH = 7.5, 5 mM $MgCl_2$, 0.2 mM EDTA, 50 mM KCl, 0.01 % Tween 20, 0.01 mg/mL BSA, and 4% glycerol by volume) (Table 7). The final concentration of the preincubation mixture was 100 nM Pol β (10X) and 50 μM inhibitor (10X). The volume of inhibitor added was unchanged across experiments that used various inhibitor concentrations. To achieve various inhibitor concentrations, the stock solution of inhibitor (1 mM) was diluted appropriately for the desired conditions (e.g. for a 10X preincubation containing 10 μM inhibitor, a 40X inhibitor solution of 400 μM was used). It should be noted that an equivalent

volume of 1:1 MeCN/H_2O (5 μL) was added to the control experiments to keep the percentage of organic solvents consistent.

Table 7. 10X Preincubation mixture for strand displacement assay(s).

	Volume added (μL)
40X 400 nM Pol β (or other polymerase)	25
40X 1 mM inhibitor (or 1:1 MeCN/H_2O for control)	5
20X reaction buffer	5
H_2O	65
Total[1]	100

[1]Preincubated at 25 °C for 20-30 min.

After 20-30 min preincubation (depending on the conditions), an aliquot (20 μL) was diluted with 1X reaction buffer (80 μL) and a 2X solution (100 μL) containing the ternary complex **B** (100 nM, 2X; Table 1) and dTTP (200 μM, 2X) (Table 8). The solution in the wells was mixed and the fluorescence measurements were collected immediately. The final mixture in the reaction well contained 10 nM Pol β, 25 μM inhibitor, 50 nM DNA, 100 μM dTTP, 1X reaction buffer, and 0.25% MeCN.

Table 8. Reaction mixture for strand displacement assay(s).

	Volume added (mL)
10X preincubation mixture	20
2X DNA and dTTP mixture	100
1X reaction buffer	80
Total	200

This method was adapted to accommodate a smaller well volume (30 μL, 384-well plates). Despite the change, enzyme and inhibitor concentrations and dilutions remained consistent (Tables 9 and 10 for volume changes).

Table 9. 10X preincubation mixture for smaller wells.

	Volume added (μL)
400 nM Pol β	12.5
1 mM inhibitor (or 1:1 MeCN/H_2O for control)	2.5
20X reaction buffer	2.5
H_2O	32.5
Total[1]	50

[1]Preincubated at 25 °C for 20-30 min.

Table 10. Reaction mixtures for smaller wells.

	Volume added (mL)
10X preincubation mixture	3
2X DNA and dTTP mixture	15
1X reaction buffer	12
Total	30

Fluorescence-based strand displacement assay for IC_{50} value at a smaller dilution

The following procedure was used to determine the IC_{50} value of **72** under previously used conditions that consist of a longer preincubation time and smaller dilution.[27]

Multiple 44X working solutions of **72** (2, 4, 10, 20, and 40 μM) were prepared in 1:1 MeCN/H_2O. A 1.1X pre-incubation mixture was prepared by mixing an 84X working solution of Pol β (2.5 μL, 200 nM) with an aliquot of **72** (5 μL, 44X, various concentrations listed above) in 1X reaction buffer (total volume: 200 μL; 50 mM HEPES buffer pH = 7.5, 5 mM $MgCl_2$, 0.2 mM EDTA, 50 mM KCl, 0.01 % Tween 20, 0.01 mg/mL BSA, and 4% glycerol by volume) (Table 11). The final preincubation mixtures contained 2.5 nM Pol β (1.1X) and inhibitor (0.05, 0.1, 0.2,

0.25, 0.5, 1 μM, 1.1X). It should be noted that an equivalent volume of 1:1 MeCN/H_2O was added to the control experiments to keep the percentage of organic solvents consistent.

Table 11. Preincubation mixture preparation.

	Volume added (μL)
84X 200 nM Pol β	2.5
44X **72** (or 1:1 MeCN/H_2O for control)	5
20X reaction buffer	10
H_2O	182.5
Total[1]	200

[1]Preincubated at 25 °C for 30 min.

After a 30 min preincubation, an aliquot (19 μL) was slightly diluted with a 20X solution (1 μL) containing the ternary complex **B** (600 nM, Table 1) and dTTP (2 mM) (Table 12). The solution was mixed and fluorescence measurements were collected immediately. The final mixture in the reaction well contained 2.38 nM Pol β, **72** (0.045, 0.09, 0.18, 0.23, 0.45, and 0.9 μM), 30 nM DNA, 100 μM dTTP, 1X reaction buffer, and 1.2% MeCN.

Table 12. Reaction mixture preparation.

	Volume added (μL)
1.1X preincubation mixture	19
20X DNA (600 nM)/dTTP (2 mM)	1
Total	20

Gel-based lyase assay (WT Pol β)

A 10X preincubation mixture was prepared by mixing a 100X working solution of Pol β (5 μL, 200 nM) with an aliquot of the inhibitor (2.5 μL, 1 mM solution in 1:1 MeCN/H_2O, 200X) in 1:1 MeCN/H_2O in 1X reaction buffer (total volume: 50 μL; 50 mM HEPES buffer pH = 7.5, 5

mM $MgCl_2$, 0.2 mM EDTA, 50 mM KCl, 0.01 % Tween 20, 0.01 mg/mL BSA, and 4% glycerol) (Table 13). The volume of inhibitor added was unchanged across experiments that used various inhibitor concentrations. To achieve various inhibitor concentrations, the stock solution of inhibitor (1 mM) was diluted appropriately for the desired conditions (e.g. for a 10X preincubation containing 10 μM inhibitor, a 40X inhibitor solution of 400 μM was used). The concentration of the preincubation mixtures was 20 nM Pol β (10X) and 50 μM inhibitor (10X). It should be noted that an equivalent volume of 1:1 MeCN/H_2O was added to the control experiments to keep the percentage of organic solvents consistent.

Table 13. 10X Preincubation mixture for gel-based lyase assay(s).

	Volume added (μL)
100X 200 nM Pol β	5
40X 1 mM inhibitor (or 1:1 MeCN/H_2O for control)	2.5
20X reaction buffer	2.5
H_2O	40
Total[1]	50

[1]Preincubated at 25 °C for 30 min.

The oligonucleotide containing 5'-dRP was generated by photolysis (350 nm, 20 min) following synthesis and purification.[29] After a 30 min preincubation, an aliquot (3 μL) was diluted with a 2X solution containing the ternary complex **A** (300 nM, 15 μL; Table 1) in 1X reaction buffer (12 μL) (Table 14). The final mixture in the reaction well contained 2 nM Pol β, 5 μM inhibitor, 150 nM DNA, 1X reaction buffer (supplemented with 0.01 mg/mL BSA, 4% glycerol), and 0.25% MeCN.

Table 14. Reaction mixtures for gel-based lyase assay(s).

	Volume added (mL)
10X preincubation mixture	3
2X DNA **A**	15
1X reaction buffer	12
Total	30

Time points of the reaction (2, 5, 10, 15, 20, and 25 min) were aliquoted and removed from the reaction mixture (4 μL). Reaction was halted by flash freezing the aliquot in dry ice. After the 25 min reaction was complete and all aliquots were frozen, the thawed mixtures were then stabilized by $NaBH_4$ (4 μL, 300 mM) for 2 hours at 4 ºC. The samples were loaded onto a 20% denaturing PAGE gel and run for 4 h at 55 watts. The gel was exposed in a radiography cassette, which was scanned using a Phosphorimager.

Gel-based lyase assay (8 kDa domain of Pol β)

A 10X preincubation mixture was prepared by mixing a 40X working solution of 8 kDa Pol β (12.5 μL, 400 nM) with an aliquot of the inhibitor (2.5 μL, 1 mM in 1:1 MeCN/H_2O, 200X) in 1:1 MeCN/H_2O in 1X reaction buffer (total volume: 50 μL; 50 mM HEPES buffer, pH = 7.5, 5 mM $MgCl_2$, 4 mM DTT, 0.01 mg/mL BSA, 4% glycerol) (Table 15). The volume of inhibitor added was unchanged across experiments that used various inhibitor concentrations. To achieve various inhibitor concentrations, the stock solution of inhibitor (1 mM) was diluted appropriately for the desired conditions (e.g. for a 10X preincubation containing 10 μM inhibitor, a 40X inhibitor solution of 400 μM was used). The concentration of the preincubation mixture was 100 nM 8 kDa domain (10X) and 50 μM inhibitor (10X).

Table 15. 10X Preincubation mixture for gel-based lyase assay(s).

	Volume added (μL)
40X 400 nM 8 kDa Pol β	12.5
200X 1 mM inhibitor (or 1:1 MeCN/H_2O for control)	2.5
20X reaction buffer	2.5
H_2O	40
Total[1]	50

[1]Preincubated at 25 ºC for 30 min.

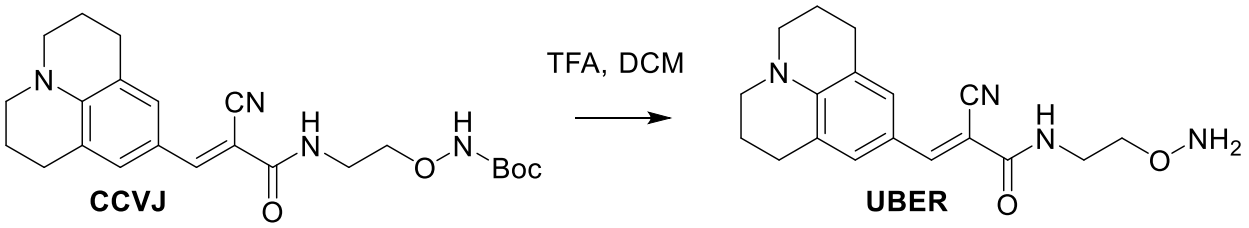
The oligonucleotide containing 5'-dRP was generated by photolysis (350 nm, 20 min) following synthesis and purification.[29] After a 30 min preincubation, an aliquot (3 μL) was diluted with a 2X solution of the ternary complex **A** (40 nM, 15 μL; Table 1) in 1X reaction buffer (12 μL) (Table 16). The final mixture in the reaction well contained 10 nM 8 kDa domain, 20 nM DNA, 1X reaction buffer.

Table 16. Reaction mixture for gel-based lyase assay(s).

	Volume added (mL)
10X preincubation mixture	3
2X DNA **A**	15
1X reaction buffer	12
Total	30

Aliquots (0, 2, 5, 10, 15, 20, and 25 min) were removed from the reaction mixture (4 μL). Reaction was halted by flash freezing the aliquot in dry ice. After 25 min, all aliquots were frozen, the time points were then stabilized by $NaBH_4$ (4 μL, 300 mM) for 2 hours at 4 ºC. The samples were loaded onto a 20% denaturing PAGE gel and run for 4 h at 55 watts. The gel was exposed in a radiography cassette, which was scanned using a Phosphorimager.

UBER probe deprotection[160]



Compound CCVJ was acquired from the Kool lab from Stanford University. CCVJ (100 nmol) was dissolved in 50 μL of a 1:1 solution of dry DCM and trifluoroacetic acid (TFA). After 4 h, the volatiles were evaporated under a stream of Argon. Remaining TFA was co-evaporated by the addition of toluene (10 L) yielding a red-orange residue. The product was confirmed by ESI-MS.

ESI-MS m/z calculated for $C_{18}H_{22}N_4O_2$ (M + H) – 327.2, 327.3 observed.

Fluorescence-based UBER lyase assay[160]

A 5X preincubation mixture was prepared by mixing a 50X working solution of Pol β (2 μL, 1 μM) with solution of inhibitor after photolysis (2 μL, 250 μM, 50X) in 1X reaction buffer (total volume: 20 μL; 50 mM HEPES buffer, pH 7.3, 5 mM $MgCl_2$, 2 mM DTT, 0.1 mg/mL BSA) (Table 17). The concentration of the preincubation mixture was 100 nM Pol β and 25 μM inhibitor (5X).

Table 17. 5X preincubation of enzyme/inhibitor for UBER assay.

	Volume added (μL)
50X 1 μM Pol β	2
50X 250 μM inhibitor (or 1:1 MeCN/H_2O for control)	2
10X reaction buffer	2
H_2O	14
Total	20

The oligonucleotide containing 5'-dRP was generated by photolysis (350 nm, 20 min) following synthesis and purification.[29] After a 20 min preincubation, DNA (1 μL, 100 μM, **A**, Table 1) was added. This resulted in a second incubation mixture containing 95 nM Pol β and 5 μM DNA. This mixture was incubated for 25 min at 25 ºC. An aliquot (4 μL) of the 5X preincubation mixture containing Pol β, inhibitor, and DNA was diluted with quenching buffer (50 mM HEPES, pH 6.5, 5 mM $MgCl_2$, 2 mM DTT, 0.1 mg/mL BSA, 2 μL). 10X Tris buffer (500 mM, pH 7.0, 1.5M NaCl) and UBER probe (10X, 50 μM in DMSO, 2 μL) were added to the 20 μL solution ([Pol β] = 20 nM, [DNA] = 1 μM, [inhibitor] = 5 μM, [UBER] = 5 μM). The reaction mixture was incubated for 4 h (Table 18). Fluorescence was measured using a Clariostar microplate reader with 5 nm slit width, 480 nm excitation, 500-560 nm emission scan.

Table 18. Reaction mixture for UBER assay.

	Volume added (μL)
5X preincubation (including **A)**	4
10X Quenching buffer	2
10X Tris buffer	2
10X 50 μM UBER probe	2
H_2O	10
Total	20

Primer Extension Assay (Pol θ)[161]

A solution of 10X **72** (5 μM or 100 μM, 2 μL) was mixed with a 40X solution of Pol θ (200 nM, 5 μL) in 1X reaction buffer (20 mM Tris•HCl pH 7.5, 100 mM NaCl, 5 mM $MnCl_2$, 0.5 mM TCEP, 10% glycerol, 0.01% NP-40, 0.1 mg/mL BSA) in a 384 microtiter well plate. This 10X preincubation mixture (50 nM Pol θ, 0.5 or 10 μM **72**) was incubated at 25 ºC for 20 min. An aliquot of the 10X preincubation mixture (2 μL) containing Pol θ and **72** was added to a new well

and mixed with 1X buffer (8 μL) and 2X cocktail solution (10 μL) containing DNA substrate **C** (1 μM), dNTPs (0.8 mM each), and 2X SYBR Gold in 1X buffer (Table 1). The final reaction mixture (20 μL) contained 5 nM Pol θ, 0.05 or 1 μM **72**, 500 nM **C**, 0.4 mM dNTPs, 1X Sybr Gold in 1X reaction buffer. Fluorescence data was collected for 80-120 min on a Varian Cary Eclipse fluorescence spectrophotometer.

Primer Extension Assays (Pol η and Pol λ)[25]

A 10X working solution of **72** (5 or 100 μM) was prepared in 1:1 MeCN/H_2O. A 10X preincubation mixture was prepared by mixing a 50X working solution of polymerase (250 nM) with an aliquot of the inhibitor (5 or 100 μM, 10X) in 1X reaction buffer (50 mM Tris•HCl, 50 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 0.1 mg/mL BSA, 10% glycerol, pH 7.5) (see Table 19 for volumes). The concentration of the preincubation mixture was 50 nM polymerase (10X) and 0, 0.5 or 10 μM **72** (1X).

Table 19. 10X preincubation mixture for primer extension assays.

	Pol η (volume added, μL)	Pol λ (volume added, μL)
50X Polymerase (250 nM)	2	3
10X **72** (5 or 100 μM) (or 1:1 MeCN/H_2O for control)	1	1.5
20X reaction buffer	1	1.5
H_2O	6	9
Total[1]	10	15

[1]Preincubated at 25 °C for 20 min.

The samples containing 50 nM polymerase and **72** (0, 0.5, or 10 μM) were preincubated at 25 °C for 20 min. After preincubation, an aliquot (2 μL, pol η or 3 μL, pol λ) was diluted with 10X ternary complex (pol η: 2 μL, 500 nM **D** or pol λ: 3 μL, 100 nM **E**), 10X dNTPs (5 mM; 2 μL, pol η or 3 μL, pol λ), and 1X reaction buffer (14 μL, pol η or 21 μL, pol λ) (see Table 20 for

volumes). While the samples were incubated at 37 ºC, aliquots (pol η: 0, 2, 5, 10, 15, 20 min, 3 μL or pol λ: 0, 5, 15, 20, 30 min, 5 μL) were removed and quenched by the addition of 95% formamide, 20 mM EDTA loading buffer (10 μL). Aliquots were heated at 95 ºC for 5 min, spun down, and loaded onto a 20% denaturing PAGE and run for 4 h at 55 watts. The gel was exposed in a radiography cassette, which was scanned using a Phosphorimager.

Table 20. Reaction mixture for primer extension assays.

	Pol η (volume added, μL)	Pol λ (volume added, μL)
10X preincubation	2	3
10X DNA	2 (500 nM **D**)	3 (100 nM **E**)
10X dNTPs (5 mM)	2	3
1X reaction buffer	14	21
Total[1]	20	30

[1]Incubated at 37 ºC, taking time points between 0-30 min.

General Procedure for Dialysis

Pol β (100 nM, total volume 200 μL) was preincubated in the absence or presence of inhibitor (e.g. 80 μM **47**) for 30 min at 25 ºC in 1X reaction buffer (50 mM HEPES buffer pH = 7.5, 5 mM $MgCl_2$, 0.2 mM EDTA, 50 mM KCl, 0.01 % Tween 20, 0.01 mg/mL BSA, and 4% glycerol). The lyase activity of each sample was immediately measured by diluting an aliquot of each sample (2.6 μL, 100 nM Pol β, ± 80 μM **47**) 10-fold by mixing with 2X ternary complex **A** (13 μL, 150 nM) in 1X reaction buffer (total volume: 26 μL) (Table 21). The final concentrations during lyase kinetics were 10 nM Pol β, 80 μM **47**, 75 nM DNA.

Table 21. Lyase reaction kinetics before dialysis.

	Volume added (μL)
10X preincubation (100 nM Pol β ± 80 μM **47)**	2.6
2X **A** (150 nM)	13
1X reaction buffer	10.4
Total	26

The lyase kinetics were carried out as previously described. Briefly, aliquots (4 μL) were removed (0, 2, 5, 10, 15, and 20 min) from the reaction mixture. The reaction was halted by flash freezing the aliquot in dry ice. After 25 min, all aliquots were stabilized by $NaBH_4$ (4 μL, 300 mM) for 1-2 h at 4 ^oC. The samples were loaded onto a 20% denaturing PAGE gel and run for 4 h at 55 watts. The gel was exposed in a radiography cassette, which was scanned using a Phosphorimager.

The remaining sample (197.4 μL) was dialyzed in a 3.50K MW cassette in reaction buffer (1 L, buffer exchanged after 12 h) containing 50 mM HEPES buffer (pH = 7.5, 5 mM $MgCl_2$, 4 mM DTT) for 24 h. The volume of the solution in the cassette was marked and no considerable volume change was observed after dialysis. The remaining lyase activity of the enzyme was measured as previously described (Table 21). Aliquots (4 μL) were removed (0, 2, 5, 10, 15, 20 min) and stabilized with $NaBH_4$ (4 μL, 300 mM) for 1-2 h at 4 ^oC. The time points were analyzed by 20% denaturing PAGE (described above).

This method was also used to analyze the effect of dialysis on the 8 kDa domain of Pol β with minor changes: (1) 1X reaction buffer contained 50 mM HEPES buffer (pH = 7.5, 5 mM $MgCl_2$, 2 mM DTT, 0.1 mg/mL BSA) and (2) the 10X 8 kDa domain preincubation concentration was increased to 200 nM, which after dilution was 20 nM. The concentration of inhibitor (80 μM) and DNA (75 nM) remained unchanged following dilution.

UPLC MS/MS Analysis of Pol β modification by covalent inhibitors

A 20X working solution of Pol β (25 μL, 20 μM, 500 pmol) was mixed with inhibitor (5 μL, 1 mM, 100X), H_2O (420 μL) and 10X reaction buffer (50 μL, 500 mM HEPES buffer, pH = 7.5, 50 mM $MgCl_2$, 20 mM DTT) to yield a mixture of 1 μM Pol β, 10 μM inhibitor, and 1X reaction buffer (50 mM HEPES buffer, pH = 7.5, 5 mM $MgCl_2$, 2 mM DTT) (Table 22).

Table 22. Preincubation for trypsin digest(s).

	Volume added (μL)
20X 20 μM Pol β	25
100X 1 mM inhibitor (or 1:1 MeCN/H_2O for control)	5
10X reaction buffer	50
H_2O	420
Total[1]	500

[1]Preincubated at 25 °C for 30 min.

After 30 min incubation, the mixture (500 μL, 1 μM Pol β) was concentrated to 50 μL (10 μM Pol β) with a 3K Amicon filter. Digestion buffer (25 μL, 500 mM Tris•HCl pH 8.0), 10X trypsin (25 μL, 400 μM), and H_2O (150 μL) were added to yield a final mixture of 2 μM Pol β and 40 μM trypsin (1:20 ratio) in 1X digestion buffer (total volume, 250 μL, 50 mM Tris•HCl pH 8.0) (Table 23).

Table 23. Trypsin digestion sample.

	Volume added (μL)
10 μM Pol β (preincubated with inhibitor and concentrated on a 3K Amicon)	50
10X digestion buffer	25
10X 400 μM trypsin	25
H_2O	150
Total	250

The digestion sample was incubated at 37 ºC overnight. A portion (100 μL) of the digestion mixture was spun down (16,000 RCF, 10 min, 4 ºC). The sample (10 μL) was injected onto and analyzed by UPLC-MS/MS using an ACQUITY UPLC HSS T3 Column (100 Å, 1.8 μm, 2.1 mm x 100 mm). The flow rate was 0.3 mL/min running a gradient from 85:5:10 water: acetonitrile: 1% formic acid to 50:40:10 water: acetonitrile: 1% formic acid over 35 min. Analysis was conducted using BioPharmaLynx with tolerance set to 30 ppm and allowing for 4 missed cleavages.

The method described above was also used to analyze the 8 kDa domain of Pol β with minor changes: (1) the 8 kDa domain of Pol β solution contained 1 nmol enzyme (2) the 10X reaction buffer contained 500 mM HEPES buffer (pH = 7.5, 50 mM $MgCl_2$, 20 mM DTT, 1 mg/mL BSA).

This method was also adapted to multiple inhibitors by varying the concentration of inhibitor during incubation. The following concentrations were used for each inhibitor analyzed: [**47**] = 10 μM, [**2**] = 5 μM, [**72**] = 300 nM. For analysis of **2** and **72**, the trypsin digestion buffer was adjusted to pH 7.1 because the adduct was reversible at pH 8.0 (Section 3.4.2.1).

Fluorescence Anisotropy[135]

Anisotropy measurements were conducted using a solution of dichloro-diphenyl-fluorescein-labeled ternary complex **F** (2.5 nM, Table 1) and Pol β (varying concentrations) in reaction buffer (50 mM HEPES, pH 7.5, 20 mM KCl, 1 mM EDTA, and 1 mM β-mercaptoethanol). Samples also contained 10% storage buffer (20 mM Tris•HCl, pH 7, 300 mM NaCl, 10% glycerol, 5 mM BME) by volume.

In a typical experiment, a sample (300 μL) was prepared by mixing Pol β (30 μL, 1 μM) in storage buffer with 10X reaction buffer (30 μL), **F** (30 μL), a solution 50% MeCN in H_2O containing or lacking **72** (200 μM, 3 μL) and H_2O (207 μL). These samples, termed solutions **A**

and **A'** (**A** did not contain **72** and **A'** contained 2 μM **68**), contained 250 pM ternary complex **F**, 100 nM Pol β, ± 2 μM **72**. Samples containing various concentrations of Pol β were prepared by serial dilution with solution **B** and **B'**. Solution **B** (10 mL) was prepared by mixing H_2O (7.85 mL), with 10X reaction buffer (1 mL), 10X storage buffer (1 mL), **F** (50 nM, 50 μL), and a solution of 50% MeCN in H_2O containing or lacking **72** (200 μM, 100 μL). Similarly, solution **B** did not contain **72** and was used exclusively to dilute solution **A**, whereas solution **B'** contained 2 μM **68** and was used to dilute solution **A'.** By mixing equal volumes of **A** or **A'** (150 μL) with **B** or **B'** (150 μL) respectively, the concentration of Pol β decreased to 50 nM, while the concentration of DNA and **72** remain unchanged. An aliquot (150 μL) of this new solution was then mixed with solution **B** or **B'** (150 μL) to prepare a new solution containing 25 nM Pol β. Serial dilutions were repeated such that samples contained Pol β concentrations of 100 nM, 50 nM, 25 nM, 12.5 nM, 6.25 nM, 3.13 nM, 1.56 nM, 0.78 nM, 0.39 nM, and 0.2 nM.

Samples were incubated at 25 ºC for 1 h and fluorescence anisotropy (A) was measured using a portion (125 μL) of each sample with a PMT voltage of 800 mV, 8 nm slit width, 535 nm excitation and 556 nm emission. Fluorescence anisotropy was measured for **F** in the absence of enzyme (A_0), and the change in anisotropy ($A-A_0$) was calculated for each sample and plotted against the concentration of Pol β. Each fluorescence anisotropy measurement was collected in triplicate.

This method was also used for the 31 kDa domain of Pol β and 8 kDa domain of Pol β. This method was adapted to analyze other inhibitors at various concentrations. The following concentrations were used: [**2**] = 30, 60 μM, [**47**] = 50 μM.

Growing conditions for different cell lines

HeLa cells were grown in DMEM with high glucose supplemented with 9% FBS and 1% antibiotic antimycotic solution (penicillin, streptomycin, and amphotericin B).

MCF-7 and MDA-MB-436 cells were grown in RPMI 1640 Medium with GlutaMAX supplement, supplemented with 0.05% insulin, 9% FBS, and 1% antibiotic antimycotic solution.

Cell viability assays

Approximately 2 x 10^5 HeLa cells were plated in each well of a 6-well culture plate (well size; 35 mm x 18 mm) in DMEM containing 10% FBS (1 mL) and kept in a humidified incubator at 37 ºC with 5% CO_2 for 20 h. A stock solution of treatments (100X **pro-72**, 10 μL; in 50% MeCN in H_2O and/or 100X MMS (20 mM), 10 μL; in DMEM-FBS medium) was diluted in culture medium (1 mL) to the appropriate concentration (**pro-72**: 0, 5, 25, 50, 100 μM; MMS: 0.2 mM) and then added to the plates keeping the quantity of the MeCN constant at 0.5% for all tests.

Cells were incubated for 1 or 2 h in a humidified incubator at 37 ºC with 5% CO_2. The medium was removed and washed with 1X PBS (2 x 1 mL). The cells were trypsinized with 0.25 w/v Trypsin-EDTA (1 mL in each well, 5 min incubation at 37 ºC), washed with DMEM-FBS (10 mL) to quench the trypsin cleavage, and spun down (3,000 RCF x 5 min). The medium was removed, and the cells were resuspended with fresh DMEM-FBS (10 mL). A portion (10 μL) of the cell suspension was thoroughly mixed with 10 μL of 0.4% solution of trypan blue in PBS (pH 7.2), and placed on a counting slide (BIO-RAD) to count the % of live cells using TC20 automated cell counter (BIO-RAD). A control experiment without treatment was carried out in parallel. All experiments were carried out in triplicate.

Clonogenic assay for cell survival

Approximately 2 x 10^5 HeLa cells were plated in each well of a 6-well culture plate (well size; 35 mm x 18 mm) in DMEM containing 10% FBS (1 mL) and kept in a humidified incubator at 37 ^oC with 5% CO_2. After overnight incubation, cells were subjected to either the vector (50% MeCN in H_2O) or treatment (100X **pro-72**, 10 μL; in 50% MeCN in H_2O and/or 100X DNA damaging agent (e.g. 20 mM MMS, 200 μM BLM; 10 μL; in DMEM-FBS medium). For alkylation experiments, cells were incubated with MMS (0 or 0.2 mM), with or without **pro-72** (5, 25 μM) at 37 ^oC with 5% CO_2 for 1 or 2 h. After treatment, the medium was removed, and the cells were washed with 1X PBS (2 x 1 mL). The cells were trypsinized with 0.25 w/v Trypsin-EDTA (1 mL in each well, 5 min incubation at 37 ^oC), washed with DMEM-FBS (10 mL) to quench the trypsin cleavage, and spun down (3,000 RCF x 5 min). The medium was removed, and the cells were resuspended with fresh DMEM-FBS (10 mL). The single cell suspensions were collected and counted using a TC20 Automated Cell Counter (BIO-RAD).

Stock solutions of single cell suspensions were prepared for all untreated and treated cells. For example, 100 cells/mL stock solution of untreated cells were prepared; 500 cells/mL stock solutions of treated cells were prepared. The concentration of stock solution for each sample was determined based on expected toxicity of the treatment (i.e. higher concentrations for more toxic conditions). The appropriate number of cells for each experiment were seeded in each well of a 6-well plate (well size; 35 mm x 18 mm) in 3 mL of DMEM-FBS medium. The cells were grown in a humidified incubator at 37 ^oC with 5% CO_2 for 14 days. No significant change in media volume was observed after two weeks due to evaporation. After 14 days, the growth medium was discarded, and the attached cells were treated with 0.2% w/v crystal violet solution. The excess

dye was washed with water. The plates were dried and scanned with an HP Scanjet 3970 and colonies were counted using ImageJ (FIJI).

Plating efficiencies (PE) and survival fractions (SF) were calculated as follows: PE = number of colonies/number of cells seeded; SF = $PE/PE_{control}$.

Pol β knock down by siRNA[179]

MDA-MB-436 cells (8 x 10^6) were seeded with RPMI 1640 medium into a 100 mm plate one day prior to transfection so cells would be adherent and ~80% confluent on the first day of transfection, determined by manual inspection. Transfection was completed as follows:

1. Lipofectamine RNAiMAX (80 μL) was diluted in Opti-MEM reduced serum medium (1 mL).
2. In a separate tube, siRNA (either siNT or siPolβ, 10 μM, 80 μL, 800 pmol) was diluted in Opti-MEM (500 μL).
3. The Lipofectamine solution was split (2 x 500 μL aliquots) and mixed with either siNT or siPolβ solution to form the lipid-RNA complex at 25 ºC for 5 min.
4. The RPMI 1640 medium was removed and replaced with Opti-MEM medium (9 mL). After 5 min incubation, the lipid-RNA solution (1 mL) was added to the cells.

The following day, steps 1-3 were repeated and the resulting 2 x 500 μL solutions were added to the corresponding cells (already in OPTI-MEM) incubating at 37 ºC, 5% CO_2. Cells were transfected with siRNA at 37 ºC, 5% CO_2 for 48 h after the second addition of siRNA. Cells were harvested and reseeded into 6-well plates for a clonogenic assay (described previously). A portion of the transfected cells (2 x 10^6) were used to determine knock down efficiency by Western blot.

This method was also used for MCF-7 cells.

Western Blot Analysis

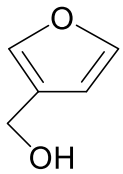
Cells were trypsinized with 0.25 w/v Trypsin-EDTA (1 mL in each well, 5 min incubation at 37 ºC), washed with RPMI 1640-FBS (10 mL) to quench the trypsin cleavage, and spun down (3,000 RCF x 5 min). The medium was removed, and cells were washed twice with cold 1X PBS buffer (2 x 3000 RCF, 4 ºC, 5 min). The cell pellet was resuspended in nuclear fractionation lysis buffer (1 mL/10^7 cells; 20 mM HEPES pH 7.5, 0.25 M sucrose, 1 mM PMSF, 3 mM $MgCl_2$, 0.2% NP-40, with protease inhibitor). Cells were incubated in lysis buffer on ice for 15 min with occasional mixing by inverting the tube back and forth, after which they were spun down (3,000 RCF, 4 ºC, 15 min). The supernatant was discarded. The cell pellet was resuspended and incubated in RIPA lysis buffer (0.5 mL/10^7 cells; 10 mM Tris pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 1% sodium deoxycholate, 0.1% SDS, 140 mM NaCl) on ice for 20 min with occasional mixing by inverting the tube back and forth. The cells were spun down (16,000 RCF, 4 ºC, 10 min). The supernatant (i.e. cell lysate) was transferred to a new tube. The concentration of proteins in cell lysate were determined using a Bradford assay (described previously). Precision Plus Protein WesternC standard (5 μL) protein ladder and Cell lysates (20 μg of protein) were separated using SDS-PAGE. Two of each sample was loaded onto the gel to visualize the amount of Pol β present and an internal standard (i.e. α-H3). The H3 antibody was used to quantify the relative amounts of H3 in the control and knock down cells to account for any variation in protein(s) loaded onto the gel. The relative amounts of H3 were used to normalize the knock down of Pol β expression observed.

Proteins were transferred from the gel to a nitrocellulose membrane (25 V, 1 A, 30 min) using a BioRad Trans-Blot Turbo Transfer system. The membrane was stained with ponceau red stain for 5 min then washed with H_2O until the background is removed. The membrane was cut

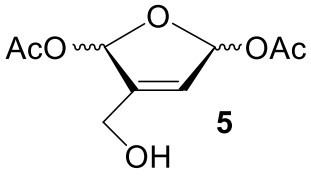
and separated so one set of samples could be incubated with the Pol β antibody and the other set with H3 antibody. The membranes were blocked with 3% BSA in 1X Tris-buffered saline tween (TBST; 20 mM Tris, 150 mM NaCl, 0.1% Tween) buffer for 30 min at 25 ºC while shaking. The membranes were incubated with the primary antibody solution (Recombinant Anti-DNA Polymerase beta antibody, 1:1000 Ab in 3% BSA in TBST; Recombinant Anti-Histone H3 antibody, 1:1000 Ab in TBST) at 4 ºC overnight. The membranes were washed (3 x 10 s TBST, 5 mL; 3 x 5 min TBST, 5 mL). In some cases, the antibody solution (α-H3) was recovered and reused. The recovered antibody solution did not contain 3% BSA, was supplemented with 5% azide after initial use, and was stored at 4 ºC. The first three rapid washes were to remove any remaining antibody solution and the latter three washes were longer to remove non-selective antibody binding. The membranes were incubated with secondary antibody solution (Rabbit anti-goat antibody, 5 mL, 1:5000 Ab in 3% BSA in TBST with StrepTactin-HRP conjugate, 0.5 μL) at 25 ºC for 1 h. The membranes were washed (3 x 10 s TBST, 5 mL; 3 x 5 min TBST, 5 mL). Membranes were kept in Tris-buffered saline (TBS; 20 mM Tris, 150 mM NaCl) buffer until incubated with developing reagent for 5 min and scanned using a Typhoon 9410 equipped with chemiluminescence imaging (Medium quality, 600 PVT, 200 pixels).

All R_f values provided in TLC information correspond to the product unless otherwise explicitly stated.

Preparation of 5[26,185]



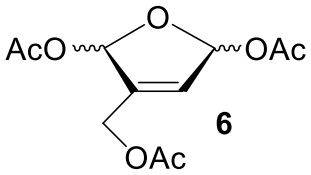
Commercially available 3-furaldehyde (4.38 mL, 5 g, 52.4 mmol, 1 eq) was added to a mixture of 10:1 Et_2O/H_2O (59 mL). The mixture was cooled to 0 ºC and stirred. Small aliquots of $NaBH_4$ (3.27 g, 86.4 mmol, 6.65 eq) were added to the mixture carefully over a period of 15 min. The reaction was stirred in air for 30 min. After 30 min, TLC (4:1 Hex/EtOAc, R_f = 0.3) showed the complete conversion of the starting material (R_f = 0.5, UV active, did not stain with PAA) to a slightly more polar spot that was not UV active and stained dark purple with PAA. The reaction was quenched with water until there was no more bubbling. The ether layer was washed with water (2 x 50 mL) and the combined aqueous layers were extracted with ether (6 x 40 mL). The final organic layer was washed with brine (1 x 50 mL), dried with $MgSO_4$, and concentrated under vacuum at 0 ºC to yield 4.714 g (91.7%) of a yellow liquid. The product was volatile, so the rotary evaporatory water bath was cooled to 0 ºC during concentration. NMR data showed the crude product was pure. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 (s, 2H), 6.38 (s, 1H), 4.46 (d, *J* = 4.9 Hz, 2H), 2.92 (d, *J* = 4.9 Hz, 1H).



5

Without purification, the intermediate (474.6 mg, 4.84 mmol, 1 eq) was combined with $Pb(OAc)_4$ (3.22 g, 7.22 mmol, 1.5 eq). Glacial acetic acid (12.5 mL) was added to the flask, which was then flushed with Argon. The reaction was stirred at 25 ºC for 21 h. When the reaction was complete by TLC (3:7 EtOAc/DCM, R_f = 0.4, stained with PAA), AcOH was removed via vacuum. Ether (50 mL) was added to the resulting residue and the precipitate was triturated with ether and removed. The filtrate was concentrated under vacuum and purified by column chromatography (7:3 Hexanes/EtOAc) to give 518 mg (58%) of a 2:1 mixture of diastereomers (**5**). ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.90 (dd, *J* = 1.0, 4.7 Hz, 1H), 6.69 (dd, *J* = 1.0, 4.7 Hz, 1H), 6.15 – 6.05 (m, 1H), 4.30 (q, *J* = 1.0 Hz, 2H), 2.19 – 2.07 (m, 6H).

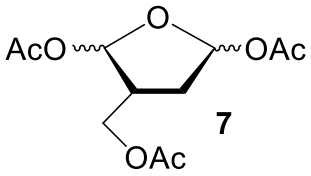
Preparation of 6



6

Compound **5** (150 mg, 0.685 mmol, 1 eq, 500 mM) and DMAP (8.6 mg, 0.07 mmol, 0.1 eq) were added to a flask. The contents were flushed with Argon and dissolved in DCM (1.25 mL). The mixture was cooled to 0 ºC. Acetic anhydride (0.33 mL, 3.5 mmol, 5 eq) was added slowly to the flask via syringe. Pyridine (0.68 mL, 8.56 mmol, 12.5 eq) was added dropwise to the solution via syringe. After 2 h, the reaction was confirmed complete by TLC (6:4 EtOAc/DCM, R_f = 0.6, stained with PAA) and was quenched with sat. $NaHCO_3$ until the pH was neutral. The mixture was diluted with EtOAc (5 mL) and washed with water (2 x 5 mL). The aqueous layer was extracted with EtOAc (2 x 10 mL), the combined organic layers were washed with brine (1 x 30 mL), dried over Na_2SO_4, and concentrated under vacuum. The crude residue was purified by column chromatography (5:1 DCM/EtOAc) to yield 167 mg (85%) of **6**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.83 (s, 1H), 6.70 – 6.56 (m, 1H), 6.22 – 5.89 (m, 1H), 4.83 – 4.57 (m, 2H), 2.11 – 1.96 (m, 9H).

Preparation of 7

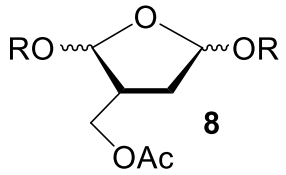


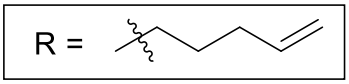
7

Compound **6** (150 mg, 0.69 mmol, 1 eq) was dissolved in EtOAc (17 mL). Rhodium on alumina catalyst (75.5 mg) was added to the pressure bottle equipped with a regulator. The vial was pressurized with H_2 to 70 psi, purged three times, and stirred at 25 ºC for 2-4 h. After venting the pressure bottle, TLC (1:1 Hex/EtOAc, R_f = 0.4, stained with PAA) showed the starting material was no longer present. When complete, the reaction mixture was passed through celite to remove the Rh catalyst. The filtrate was concentrated under vacuum to give 127.5 mg (85%) of pale, yellow compound **7**. No purification was needed. ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.25 – 6.15 (m, 2H), 4.06 (m, 1H), 4.03 – 3.91 (m, 1H), 2.71 – 2.56 (m, 1H), 2.38 (m, 1H), 2.04 – 1.88 (m, 9H), 1.83 – 1.67 (m, 1H).

^{13}C NMR ($CDCl_3$) 170.7, 169.9, 169.5, 98.5, 96.4, 62.1, 41.6, 32.7, 21.5, 20.9, 20.7.

Preparation of 8

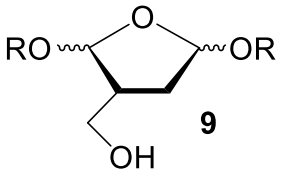




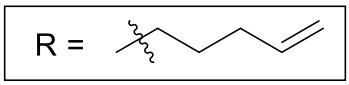
BF_3•etherate was distilled from CaH_2 under vacuum and kept under Argon. Compound **7** (750 mg, 2.88 mmol, 1 eq, 160 mM) was dissolved in DCM (18 mL) and cooled to 0 ºC. 4-Pentenol (1.86 mL, 1.49 g, 17.3 mmol, 6 eq) was added to the reaction. BF_3•etherate (8.14 mL, 1.67 M) was slowly added to the solution over a period of 15-20 min (until diluted to 28 mL, 600 mM). After 30 min, the reaction was incomplete when analyzed by TLC (7:3 Hex/Et_2O, stained with PAA) but degradation products began to appear, so the reaction was quenched with sat. $NaHCO_3$ (5 mL), diluted with DCM (20 mL), and washed with sat. $NaHCO_3$ (1 x 15 mL). The aqueous layer was extracted with DCM (4 x 20 mL). The organic layer was washed with water (1 x 30 mL), brine (1 x 30 mL), and dried over Na_2SO_4. The residue was concentrated under vacuum and purified by column chromatography (8:2 hex/EtOAc) to give 228.6 mg (25.4%) of **8a** and 519 mg (57.7%) of **8b** (totaling 83%). **8a** ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.82 – 5.65 (m, 2H), 5.21 – 4.84 (m, 6H), 4.24 – 3.99 (m, 2H), 3.75 – 3.57 (m, 2H), 3.45 – 3.22 (m, 2H), 2.61 (m, 1H), 2.42 – 2.17 (m, 1H), 2.08 – 2.01 (m, 4H), 1.99 (d, *J* = 9.5 Hz, 3H), 1.93 – 1.74 (m, 1H), 1.67 – 1.50 (m, 4H). **8b** ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.91 – 5.68 (m, 2H), 5.10 – 4.83 (m, 6H), 4.15 – 3.88 (m, 2H), 3.73 (m, 2H), 3.55 – 3.32 (m, 2H), 2.65 (ddd, *J* = 5.5, 7.4, 13.5 Hz, 1H), 2.39 (s, 1H), 2.32 – 2.18 (m, 1H), 2.16 – 2.08 (m, 4H), 2.05 (s, 3H), 1.81 (ddd, *J* = 5.5, 7.4, 13.5 Hz, 1H), 1.73 – 1.59 (m, 4H).

8a and **8b** ^{13}C NMR ($CDCl_3$) 170.9, 170.8, 138.20, 138.18, 138.15, 114.8, 106.9, 106.2, 104.7, 104.2, 103.5, 103.2, 67.6, 67.5, 67.4, 67.3, 67.2, 67.0, 64.8, 64.5, 63.3, 43.8, 43.2, 40.7, 34.7, 34.2, 33.4, 30.4, 30.3, 29.9, 28.9, 28.8, 20.9, 20.8.

Preparation of 9[26]

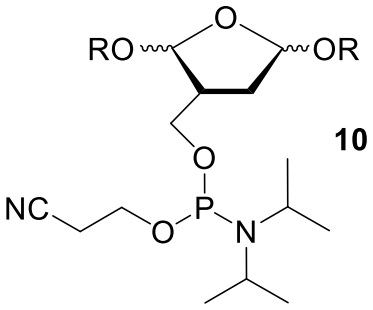


9

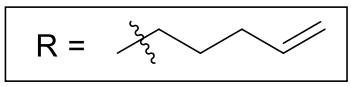


A diastereomer of **8** (82 mg, 0.26 mmol, 1 eq, 150 mM) was dissolved in MeOH (1.7 mL). Sodium methoxide stock solution (700 mM) was prepared by dissolving Na metal (122 mg, 5.3 mmol) in MeOH (7.5 mL). An aliquot of NaOMe (700 mM, 0.3 mL) was added to the reaction slowly (effectively diluting NaOMe to 100 mM). After 2 h, TLC (6:4 Hex/EtOAc, R_f = 0.2, stained with PAA) confirmed the reaction was complete by the disappearance of **8**. The reaction was quenched with a few drops of AcOH until neutral pH. The reaction was diluted with DCM (20 mL) and washed with H_2O (2 x 15 mL). The aqueous layer was extracted with DCM (4 x 20 mL) and the combined organic layers were washed with brine (1 x 30 mL) and dried over Na_2SO_4. The reaction was concentrated under vacuum to give 49.3 mg (70%) of **9**. No purification was needed. ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.91 – 5.65 (m, 2H), 5.23 – 4.88 (m, 6H), 3.93 – 3.71 (m, 2H), 3.72 – 3.61 (m, 2H), 3.49 – 3.31 (m, 2H), 2.63 – 2.46 (m, 1H), 2.38-2.23 (ddd, *J* = 5.0, 9.5, 13.3 Hz, 1H), 2.26 – 2.13 (m, 1H), 2.13 – 2.03 (m, 4H), 1.82 (ddd, *J* = 1.5, 7.8, 13.3 Hz, 1H), 1.65 (dddt, *J* = 1.5, 5.0, 7.8, 9.5 Hz, 4H).

Preparation of 10



10

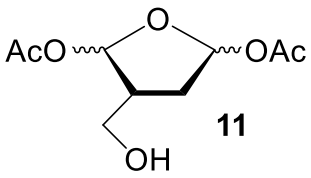


Compound **9** (58.2 mg, 0.22 mmol) was azeotropically dried with pyridine (2 x 0.5 mL). The reagent was cooled to 0 °C. DIPEA (0.18 mL, 133 mg, 0.88 mmol, 4 eq) was added to the cold starting material and the reactants were dissolved in DCM (2.1 mL, 100 mM). 2-Cyanoethyl-N, N-diisopropylchlorophosphoramidite (0.06 mL, 60.4 mg, 0.26 mmol. 1.2 eq) was added and the cold mixture stirred with periodic monitoring by TLC (7:1 Hex/EtOAc, R_f = 0.3, stained with PAA). After 2 h, TLC showed complete conversion to the phosphoramidite. The reaction was diluted with freshly distilled EtOAc (10 mL). The organic layer was washed with saturated

bicarbonate solution (2 x15 mL) and the aqueous layers were extracted with distilled EtOAc (2 x 20 mL). The combined organic layers were washed with brine (1 x 25 mL) and dried over Na_2SO_4. The organic layer was concentrated under vacuum and purified by column chromatography (7:1 distilled hexanes/distilled EtOAc) yielding 55.9 mg (54%) of **10**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.80 (dddt, *J* = 4.1, 8.2, 13.5, 16.0 Hz, 2H), 5.18 – 4.93 (m, 6H), 3.86 – 3.72 (m, 2H), 3.64 (dtd, *J* = 2.1, 6.0, 8.2 Hz, 2H), 3.59 – 3.51 (m, 2H), 3.34 (dddd, *J* = 2.1, 4.1, 9.4, 16.0 Hz, 2H), 2.71 – 2.52 (m, 2H), 2.30 (ddd, *J* = 6.0, 9.4, 13.5 Hz, 2H), 2.11 – 2.02 (m, 4H), 1.93 – 1.76 (m, 1H), 1.65 – 1.55 (m, 6H), 1.17 (m, 12H).

^{31}P NMR (400 MHz, $CDCl_3$) δ 147.87.

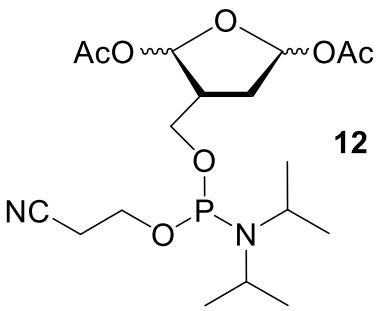
Preparation of 11[27]



11

Compound **5** (163 mg, 0.886 mmol, 1 eq) was dissolved in EtOAc (16 mL, 50 mM). Rhodium on alumina catalyst (85.2 mg) was added to the pressure bottle equipped with a regulator. The vial was pressurized with H_2 to 70 psi, purged three times, and stirred at 25 ºC for 2-4 h. After venting the pressure bottle, TLC (4:6 Hex/EtOAc, R_f = 0.3, stained with PAA) showed the starting material was no longer present. When complete, the reaction mixture was passed through celite to remove the Rh catalyst. The filtrate was concentrated under vacuum to a pale, yellow residue, which was purified by column chromatography (1:1 DCM/EtOAc) resulting in 67 mg (41%) of a mixture of diastereomers (**11**). ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.44 (s, 1H), 6.34 (d, *J* = 1.0 Hz, 1H), 3.86 – 3.75 (m, 1H), 3.69 (dt, *J* = 5.5, 10.9 Hz, 1H), 2.56 – 2.43 (m, 2H), 2.05 (s, 3H), 2.04 (s, 3H), 1.90 (t, *J* = 5.5 Hz, 1H), 1.82 (dd, *J* = 1.0, 5.5 Hz, 1H).

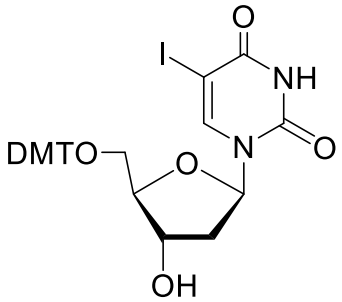
Preparation of 12[27]



Compound **11** (67 mg, 0.28 mmol) was azeotropically dried with pyridine (2 x 0.5 mL). The flask was cooled to 0 °C. DIPEA (0.2 mL, 144 mg, 1.12 mmol, 4 eq) was added to the cold starting material and the reactants were dissolved in DCM (1.2 mL, 200 mM). 2-Cyanoethyl-N, N-diisopropylchlorophosphoramidite (75 μL, 80 mg, 0.34 mmol. 1.2 eq) was added and the cold mixture stirred with periodic monitoring by TLC (7:1 Hex/EtOAc, R_f = 0.2, stained with PAA). After 2 h, TLC showed complete conversion to the phosphoramidite. The reaction was diluted with EtOAc (10 mL). The organic layer was washed with saturated bicarbonate solution (2 x15 mL) and the aqueous layers were extracted with EtOAc (2 x 20 mL). The combined organic layers were washed with brine (1 x 25 mL) and dried over Na_2SO_4. The organic layer was concentrated under vacuum and purified by column chromatography (3:1 hexanes/ EtOAc) yielding 85 mg (75%) of **12**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.31 (m, 1H), 4.28 (m, 1H), 3.78 (m, 3H), 3.55 (m, 3H), 2.79 (m, 1H), 2.60 (m, 2H), 2.44 (m, 1H), 2.04 (m, 6H), 1.81 (m, 1H), 1.14 (m, 12H).

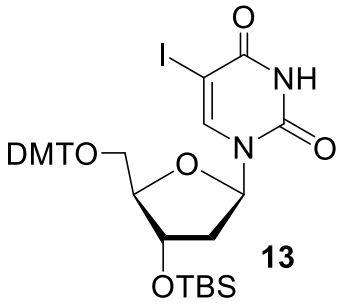
^{31}P NMR (162 MHz, CDCl3) δ 149.7, 149.5, 148.3, 148.1, 13.2.

Preparation of 13[186–188]



5-Iodo-2'-deoxyuridine (953 mg, 2.69 mmol) was azeotropically dried with pyridine (2 x 3 mL). DMT-Cl (1.63 g, 4.8 mmol, 1.8 eq) was added and reactants were dissolved in pyridine (15 mL, 180 mM) at 0 °C. The mixture turned orange and was monitored by TLC (5% MeOH in DCM, R_f = 0.7, UV active and stained orange then blue with PAA). After 2 h, side products (presumably the bis-tritylated product) were observed by TLC. The reaction was concentrated under vacuum to remove pyridine. The residue was dissolved in EtOAc (30 mL) and

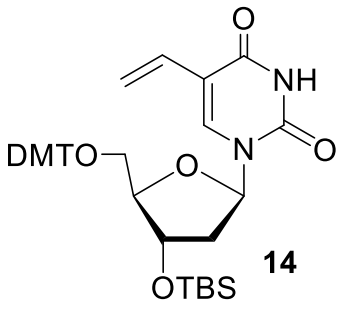
became yellow. The solution was washed with sat. sodium bicarbonate (1 x 30 mL) and H_2O (1 x 30 mL). The aqueous layer was extracted with EtOAc (2 x 30 mL) and the combined organic layers were washed with brine (1 x 50 mL), dried over Na_2SO_4, and concentrated under vacuum. The crude material was purified by column chromatography (2:1 EtOAc/hexanes) to yield 1.53g (87%) of the DMT intermediate. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.12 (s, 1H), 7.41 (m, 2H), 7.32 (dd, *J* = 2.7, 9.0 Hz, 6H), 7.24 (m, 4H) 6.84 (d, *J* = 9.0 Hz, 4H), 6.30 (dd, *J* = 5.7, 7.7 Hz, 1H), 4.53 (dt, *J* = 2.7, 5.7 Hz, 1H), 4.07 (q, *J* = 3.3 Hz, 1H), 3.79 (s, 6H), 3.42 (dd, *J* = 3.3, 10.7 Hz, 1H), 3.36 (dd, *J* = 3.3, 10.7 Hz, 1H), 2.47 (ddd, *J* = 2.7, 5.7, 13.6 Hz, 1H), 2.28 (ddd, *J* = 5.7, 7.7, 13.6 Hz, 1H).



13

The DMT intermediate (415 mg, 0.633 mmol) was azeotropically dried with pyridine (2 x 3 mL), while DMF was freshly distilled from CaH_2. Imidazole (123 mg, 1.8 mmol, 2.8 eq) was added to the flask with starting material and the flask was resealed and flushed with Argon. Reagents were dissolved in DMF (4.2 mL, 150 mM). TBDMSCl (114.1 mg, 0.76 mmol, 1.2 eq) was added and the mixture was heated to 55 ^oC with a water condenser. After 21 h, TLC (5% MeOH in DCM, R_f = 0.9, UV active and stained orange then blue with PAA) confirmed reaction completion. The reaction was cooled, diluted with EtOAc (10 mL), and washed with brine (2 x 10 mL). The aqueous layer was extracted with EtOAc (2 x 10 mL), and the combined organic layers were dried over Na_2SO_4 and concentrated under vacuum. The crude residue was purified by column chromatography (1:1 hexanes/EtOAc) to yield 386 mg (80%) of **13**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.11 (s, 1H), 8.19 (s, 1H), 7.40 (m, 2H), 7.36 – 7.29 (m, 5H), 7.20 (m, 2H), 6.84 (m, 4H), 6.27 (dd, *J* = 5.8, 7.6 Hz, 1H), 4.48 – 4.33 (m, 1H), 3.99 (q, *J* = 2.9 Hz, 1H), 3.78 (d, *J* = 0.5 Hz, 6H), 3.40 (dd, *J* = 2.9, 10.8 Hz, 1H), 3.27 (d, *J* = 10.8 Hz, 1H), 2.37 (ddd, *J* = 2.7, 5.8, 13.3 Hz, 1H), 2.18 (ddd, *J* = 5.8, 7.6, 13.3 Hz, 1H), 0.83 (s, 9H), -0.01 (d, *J* = 0.5 Hz, 6H).

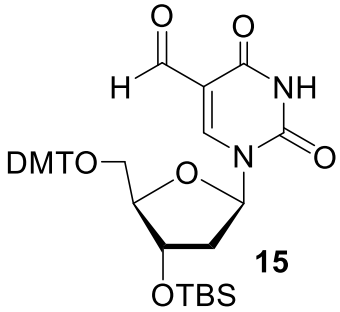
Preparation of 14[188]



14

Compound **13** (346 mg, 0.5 mmol) and $Pd(PPh_3)Cl_2$ (32.2 mg, 0.05 mmol, 0.1 eq) were flushed with Argon. Tributyl vinyl tin (0.17 mL, 184.6 mg, 0.6 mmol, 1.3 eq) was added. DMF was freshly distilled from CaH_2 and sparged with Ar for 45 minutes. DMF (0.87 mL, 150 mM) was added to the reaction mixture and the flask was heated at reflux for 2 h. When TLC (5:1 Hex/EtOAc, R_f = 0.5, product was UV active and stained orange then pink then blue with PAA) confirmed reaction was complete, the reaction was cooled and diluted with EtOAc (3 mL), washed with sat. bicarbonate solution (2 x 3 mL), H_2O (1 x 3 mL), and brine (1 x 3 mL). The aqueous layer was extracted with EtOAc (1x 10 mL) and the combined organic layers were washed with brine, dried over Na_2SO_4, and concentrated under vacuum. The crude material was purified by column chromatography (5:1 Hex/EtOAc) to yield 208 mg **14** (70%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.18 (s, 1H), 7.71 (s, 1H), 7.35 (m, 2H), 7.29 (s, 4H), 7.20 (m, 3H), 6.86 – 6.69 (m, 4H), 6.30 (d, *J* = 6.7 Hz, 1H), 5.78 (m, 2H), 4.90 (m, 1H), 4.45 (dt, *J* = 3.3, 6.7 Hz, 1H), 3.95 (d, *J* = 3.3 Hz, 1H), 3.75 (s, 6H), 3.43 (dd, *J* = 3.2, 10.7 Hz, 1H), 3.25 (dd, *J* = 3.2, 10.7 Hz, 1H), 2.33 (ddd, *J* = 3.3, 6.6, 13.3 Hz, 1H), 2.16 (dt, *J* = 6.6, 13.3 Hz, 1H), 0.80 (s, 9H), -0.04 (s, 3H), 0.00 (s, 3H).

Preparation of 15

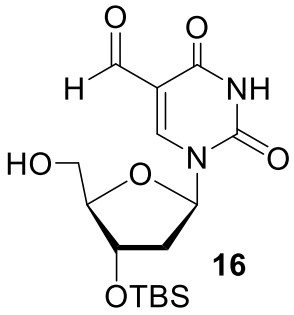


15

Compound **14** (197 mg, 0.294 mmol) and N-methylmorpholine N-oxide (69.6 mg, 0.6 mmol, 2 eq) were dissolved in acetone/H_2O/tBuOH (4:1:1) mixture (2 mL, 150 mM). When the reactants were dissolved, 4% OsO_4 in water (50 μL) was added and the mixture was stirred open to air at 25 °C for 2 h. When TLC (5% MeOH in DCM, R_f = 0.5, UV active and stained orange then blue with PAA) showed complete conversion, $NaIO_4$ (86.8 mg, 0.353 mmol, 1.2 eq) was added and the reaction stirred for one hour until complete. The reaction was diluted with

EtOAc (4 mL) and H_2O (2 mL). The aqueous layer was extracted with EtOAc (3 x 5 mL). The combined organic layers were washed with brine (1 x 15 mL), dried over Na_2SO_4, and concentrated under vacuum. The crude material was purified by column chromatography (5:1 -> 3:1 Hex/EtOAc) to yield 134 mg of **15** (68%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.71 (s, 1H), 8.49 (s, 1H), 7.34 (d, *J* = 7.1 Hz, 2H), 7.31 – 7.06 (m, 6H), 7.16 (m, 1H), 6.80 (d, *J* = 7.1 Hz, 4H), 6.14 (m, 1H), 4.27 (m, 1H), 4.04 (q, *J* = 3.8 Hz, 1H), 3.74 (s, 6H), 3.36 (dd, *J* = 3.8, 10.8 Hz, 1H), 3.29 (dd, *J* = 3.8, 10.8 Hz, 1H), 2.45 (m, 1H), 2.26 – 2.06 (m, 1H), 0.80 (s, 9H), -0.05 (s, 3H), -0.02 (s, 3H).

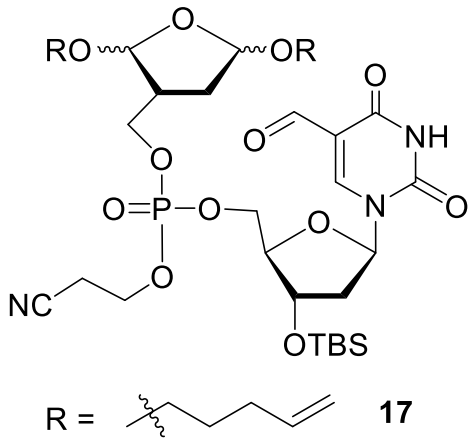
Preparation of 16



16

Aqueous trichloroacetic acid (TCA) solution (0.6 M) was prepared by dissolving TCA crystals (96.7 mg, 0.59 mmol) in distilled H_2O (1 mL). Compound **15** (207 mg, 0.31 mmol) was dissolved in DCM (0.91 mL). The TCA solution (0.6 M, 0.63 mL) was added to the reaction flask (TCA (242 mM) and **15** (200 mM)). After 20 min, TLC (5% MeOH in DCM, R_f = 0.3, UV active and stained blue with PAA) showed complete conversion. MeOH was avoided in all steps containing the aldehyde to avoid forming the acetal. When complete, the reaction was quenched with sat. bicarbonate solution until neutral pH, diluted with DCM (2 mL), and washed with bicarbonate solution (1 x 3mL). The aqueous layer was extracted with DCM (4 x 4 mL) and the combined organic layers were washed with brine (1x 15 mL), dried over Na_2SO_4, and concentrated under vacuum. Column chromatography (2:1 -> 1:1 Hex/EtOAC) yielded 102 mg (91%) of **16**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.91 (s, 1H), 9.88 (s, 1H), 8.87 (s, 1H), 6.19 (t, *J* = 6.2 Hz, 1H), 4.48 (dt, *J* = 4.2, 6.2 Hz, 1H), 4.00 – 3.90 (m, 2H), 3.78 (s, 1H), 2.37 (ddd, *J* = 4.2, 6.3, 13.5 Hz, 1H), 2.25 (dt, *J* = 6.3, 13.5 Hz, 1H), 0.85 (s, 9H), 0.05 (s, 6H).

Preparation of 17



Compounds **10** (92.1 mg, 0.2 mmol, 1 eq) and **16** (87 mg, 0.23 mmol, 1.2 eq) were azeotropically dried together with toluene (2 x 2 mL). S-Ethyl tetrazole/ACN (250 mM, 0.96 mL, 0.23 mmol, 1.2 eq) was added to the flask. After 3 h of stirring at 25 ºC, TLC (2.5:1 EtOAc/Hex, R_f = 0.46 UV active and stained with PAA) indicates that the majority of **16** was consumed. tBuOOH (500 mM, 1.17 mL, 0.6 mmol, 3 eq) was added to the flask for 15-20 min. The contents of the flask were concentrated under vacuum and purified by column chromatography (6:1 DCM/EtOAc) to yield 97.1 mg (66%) of **17**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.90 (s, 1H), 8.35 (m, 1H), 6.16 (dt, *J* = 2.9, 6.6 Hz, 1H), 6.01 (m, 1H), 5.87 – 5.67 (m, 3H), 5.16 – 4.80 (m, 6H), 4.76 (m, 1H), 4.40 (m, 1H), 4.28 (m, 4H), 3.64 (td, *J* = 2.9, 6.6 Hz, 2H), 3.50 (m, 1H), 3.34 (m, 1H), 2.76 (m, 2H), 2.36 (m, 1H), 2.14 (m, 1H), 2.04 (m, 4H), 2.00 – 1.73 (m, 1H), 1.60 (m, 5H), 0.84 (s, 9H), 0.04 (s, 6H).

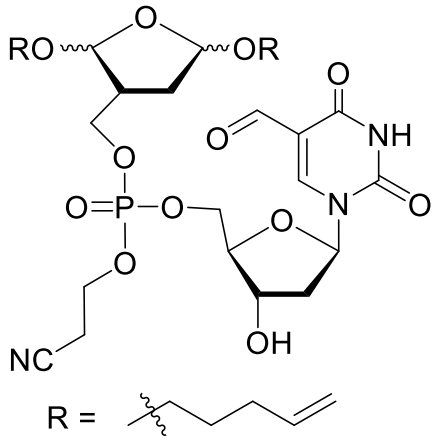
^{13}C NMR (400 MHz, $CDCl_3$) δ 186.1, 106.1, 104.1, 88.3, 86.7, 85.9, 67.9, 41.6, 30.24. 30.20, 30.15, 28.94, 28.86, 28.8, 25.7, 25.64, 25.60. 25.58, -4.7, -4.9.

^{31}P NMR (400 MHz, $CDCl_3$) δ -1.8.

IR: 3184, 3076, 2931, 2858, 2254, 1701, 1601, 1471, 1375, 1275.

ESI-TOF m/z calculated for $C_{34}H_{54}N_3O_{12}PSi$ (M + Na) − 778.3111, 778.3085 observed).

Preparation of 18



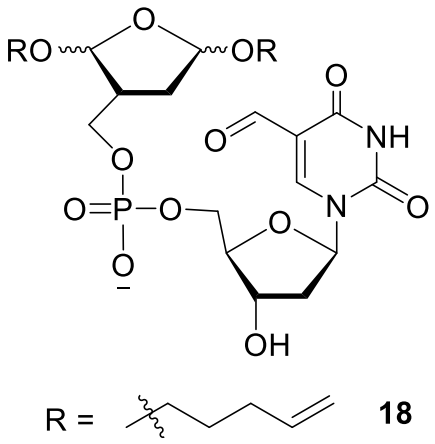
Compound **17** (88.8 mg, 0.12 mmol) was dissolved in THF (1.2 mL, 100 mM). TEA•3HF (590 µL, 0.6 mmol, 5eq) was added to the flask and the reaction was stirred overnight. After confirmed complete by TLC (5% MeOH in DCM, R_f = 0.3, UV active and stained with PAA), the reaction was diluted with DCM (1 mL) and washed with H_2O (1 x 2 mL), bicarb (1 x 2 mL), and brine (1 x 2 mL). The combined organic layers were dried over Na_2SO_4 and concentrated under vacuum. Column chromatography (3% MeOH in DCM) provided 60.5 mg (81%) of the intermediate. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.87 (s, 1H), 8.37 (m, 1H), 6.18 (dt, *J* = 3.5, 6.3 Hz, 1H), 5.78 (m, 2H), 5.58 (m, 1H), 5.19 – 4.80 (m, 6H), 4.49 (s, 1H), 4.34 (m, 5H), 4.16 (dd, *J* = 3.5, 6.3 Hz, 2H), 4.00 (m, 1H), 3.76 – 3.60 (m, 2H), 3.55 (dd, *J* = 2.5, 6.3 Hz, 1H), 3.36 (m, 1H), 2.80 (t, *J* = 3.6 Hz, 2H), 2.50 (m, 1H), 2.23 (m, 2H), 2.07 (dd, *J* = 2.5, 3.6 Hz, 4H), 1.82 – 1.49 (m, 4H), 1.20 (m, 1H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 149.4, 138.12, 138.08, 138.07, 137.97, 137.9, 114.7, 114.6, 111.2, 105.2, 67.3, 67.1, 30.2, 30.10, 30.08, 30.0, 28.8, 28.7, 19.7, 19.6.

^{31}P NMR (400 MHz, $CDCl_3$) δ -1.60, -1.61, -1.62, -1.64, -1.70, -1.74, -1.78, -1.80, -1.82, -1.85, -1.90, -1.94.

IR: 3410, 3075, 2939, 2255, 1693, 1601, 1471, 1274, 1027.

ESI-TOF m/z calculated for $C_{28}H_{40}N_3O_{12}P$ (M + Na) − 664.2247, 664.2228 observed).



R = **18**

The intermediate (141.1 mg, 0.22 mmol) was dissolved in freshly distilled DMF (550 μL) and DIPEA (550 μL, 50% final solution). The reaction flask was heated to 50 ºC for 25 h until complete as seen by TLC (20% MeOH in DCM, R_f = 0.3, UV active and stained with PAA). The reaction was cooled and concentrated under vacuum. Column chromatography (15% MeOH in DCM) provided 64.6 mg (50%) of **18**. ^{1}H NMR (400 MHz, CD_3OD) δ 9.70 (m, 1H), 8.65 (m, 1H), 6.16 (dd, *J* = 4.6, 7.8 Hz, 1H), 5.74 (m, 2H), 5.53 (m, 1H), 5.15 – 4.81 (m, 6H), 4.40 (m, 2H), 4.03 (m, 2H), 3.97 (m, 1H), 3.81 (m, 1H), 3.61 (m, 2H), 3.50 (m, 2H), 3.28 (m, 2H), 3.23 (m, 1H), 2.35 (m, 1H), 2.19 (m, 1H), 2.02 (t, *J* = 7.8 Hz, 4H), 1.67 – 1.46 (m, 4H).

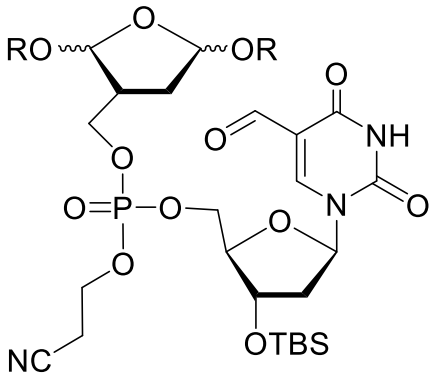
^{13}C NMR (400 mHz, CD_3OD) δ 143.5, 97.1, 91.1, 83.7, 72.0, 59.8, 59.7, 58.6, 58.5, 54.8, 41.4.

^{31}P NMR (400 MHz, CD_3OD) δ -0.1.

IR: 3266, 2940, 2361, 2340, 1701, 1472, 1228, 1094, 1057.

ESI-TOF m/z calculated for $C_{25}H_{36}N_2O_{12}P^-$ (M + H + Na) − 611.1987, 611.1959 observed).

Preparation of 19



R = -Ac **19**

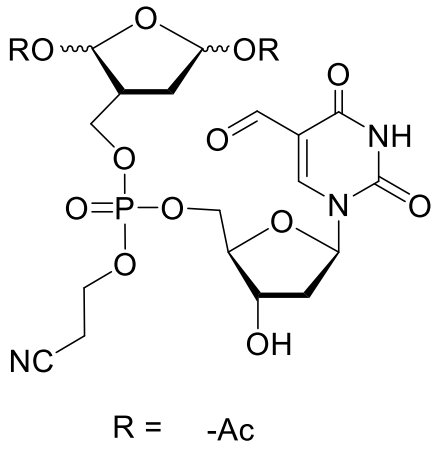
Compounds **12** (140 mg, 0.55 mmol, 1.2 eq) and **16** (242 mg, 0.42 mmol, 1 eq) were azeotropically dried together with degassed toluene (2 x 2 mL). S-Ethyl tetrazole/MeCN (250 mM, 2.2 mL, 0.55 mmol, 1.2 eq) was added to the flask. After 3.5 h of stirring at 25 ºC, TLC (2.5:1 EtOAc/Hex, R_f = 0.6, UV active and stained with PAA) indicated the consumption of **16**. tBuOOH (500 mM, 0.3 mL, 1.26 mmol, 3 eq) was added for 15-20 min. The contents of the flask were concentrated under vacuum and column

chromatography (2:1 EtOAc/Hex -> 3:1 EtOAc/Hex) provided 151 mg (50%) of **19.** ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.91 (s, 1H), 8.36 (s, 1H), 6.33 (s, 1H), 6.17 (s, 1H), 4.41 (s, 1H), 4.29 (m, 6H), 4.07 (m, 4H), 2.77 (s, 3H), 2.40 (s, 1H), 2.14 (m, 1H), 2.02 (s, 3H), 1.99 (s, 3H), 1.83 (s, 1H), 1.20 (s, 9H), 0.04 (s, 6H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 196.2, 171.2, 169.9, 150.9, 149.2, 136.5, 98.4, 96.0, 86.7, 85.8, 77.4, 77.1, 76.9, 62.7, 60.4, 56.0, 41.5, 39.2, 34.6, 32.6, 31.5, 29.6, 29.0, 25.6, 25.2, 22.6, 21.1, 20.1, 19.7, 17.8, 14.5, 14.1, 12.3, 11.4, -4.7, -4.9.

^{31}P NMR (400 MHz, $CDCl_3$) δ 8.91, 8.6, -1.6, -1.7, -1.9, -2.0.

Preparation of 20

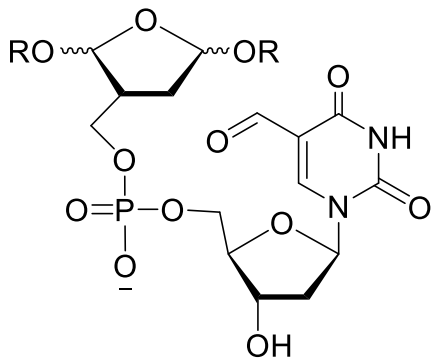


Compound **19** (130 mg, 0.185 mmol) was dissolved in THF (1.85 mL, 100 mM). TEA•3HF (150 µL, 0.925 mmol, 5 eq) was added and reaction stirred overnight. After confirmed complete by TLC (5% MeOH in DCM, R_f = 0.25, UV active and stained with PAA), the reaction was diluted with DCM (2 mL) and washed with H_2O (1 x 3 mL), sat. sodium bicarbonate (1 x 3 mL), and brine (1 x 3 mL). The combined organic layers were dried over Na_2SO_4 and concentrated under vacuum. Column chromatography (3% MeOH in DCM) yielded 109 mg (100%) of the desilylated intermediate. ^{1}H NMR (400 MHz, CD_3CN) δ 9.80 (s, 1H), 8.48 (s, 1H), 6.33 (s, 1H), 6.29 (s, 2H), 4.56 (s, 1H), 4.39 (s, 4H), 4.30 (m, 3H), 4.11 (d, *J* = 2.8, 2H), 4.02 (s, 2H), 2.77 (dd, *J* = 1.8, 2.8 Hz, 2H), 2.47 (d, *J* = 1.8 Hz, 1H), 2.31 (d, *J* = 5.2 Hz, 1H), 2.17 (d, *J* = 5.2 Hz, 1H), 2.05 – 1.90 (m, 6H).

^{13}C NMR (400 MHz, CD_3CN) δ 205.6, 186.1, 170.5, 170.1, 166.4, 161.5, 149.5, 146.2, 111.3, 105.0, 98.4, 86.8, 59.7, 54.2, 46.2, 40.4, 29.6, 29.4, 29.2, 29.0, 28.8, 28.7, 28.5, 20.4, 20.3, 20.0,

19.2, 13.6, 8.0.

^{31}P NMR (400 MHz, CD_3CN) δ -1.72, -1.74.



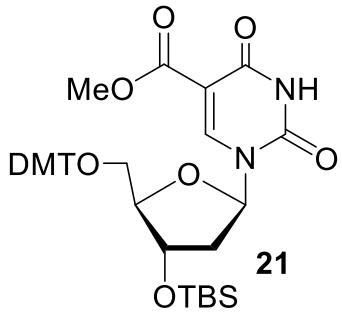
R = -Ac **20**

The intermediate (109 mg, 0.185 mmol) was dissolved in DMF (460 µL) and DIPEA (460 µL, 50% final solution). The reaction was heated to 50 °C for 25 h until complete as seen by TLC (20% MeOH in DCM, R_f = 0.15, UV active and stained with PAA). The reaction was cooled and concentrated under vacuum. Column chromatography (15% MeOH in DCM) provided 92 mg of **20** (94%). ^{1}H NMR (400 MHz, CD_3CN) δ 9.76 (s, 1H), 8.55 (s, 1H), 6.22 (dd, *J* = 5.0, 6.2 Hz, 1H), 6.16 (d, *J* = 5.0, Hz, 1H), 6.08 (t, *J* = 6.2 Hz, 1H), 4.37 (s, 2H), 4.01 (m, 2H), 3.95 – 3.85 (m, 2H), 3.57 (m, 2H), 2.71 (m, 1H), 2.44 (m, 1H), 2.34 (m, 1H), 2.21 (m, 1H), 1.98 (s, 6H).

^{13}C NMR (400 MHz, CD_3CN) δ 187.7, 171.7, 171.0, 170.9, 150.5, 118.3, 112.3, 99.4, 97.4, 87.5, 71.2, 65.4, 60.9, 55.2, 54.3, 49.8, 47.4, 42.7, 30.8, 21.3, 21.0, 19.2, 18.0, 14.4, 12.8, 1.8, 1.6, 1.40, 1.38, 1.36, 1.19, 1.17, 1.1, 1.0, 0.8, 0.7, 0.5.

^{31}P NMR (400 MHz, CD_3CN) δ -2.3.

Preparation of 21

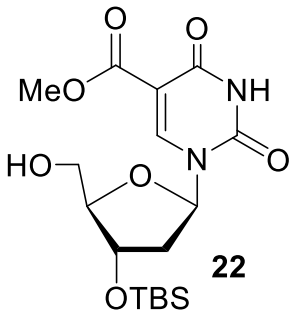


21

Compound **13** (1.23 g, 1.5 mmol) and $Pd(CH_3CN)Cl_2$ (45.4 mg, 0.002 mmol, 0.1 eq) were added to a pressure flask, sealed with a septum, and dried under vacuum for 15 min. Under Argon, triethylamine (0.8 mL, 5.4 mmol, 3.6 eq) and freshly distilled MeOH (9 mL, 150 mM) were added to the flask. The pressure vessel was purged with CO three times and held at 24 psi. The mixture turned black while purging with CO. The vessel was

heated at 50 °C for 3 h until the reaction was complete by TLC (10% MeOH in DCM, R_f = 0.5, UV active and stained by PAA). The crude mixture was passed through a Celite column and concentrated under vacuum. The crude mixture was purified by column chromatography (2:1 Hex/EtOAc -> 1:1 Hex/EtOAc) to yield 811mg of **21** (77%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.77 (s, 1H), 8.01 (s, 1H), 7.42 (d, *J* = 7.2 Hz, 2H), 7.32 (dd, *J* = 6.4, 8.9 Hz, 5H), 7.22 (m, 2H), 6.82 (dd, *J* = 3.3, 8.9 Hz, 4H), 6.23 (d, *J* = 6.0 Hz, 1H), 4.31 (td, *J* = 3.3, 6.4 Hz, 1H), 3.99 (d, *J* = 3.2 Hz, 1H), 3.79 (d, *J* = 0.5 Hz, 6H), 3.55 (dd, *J* = 3.2, 10.9 Hz, 1H), 3.21 (s, 3H), 2.41 (ddd, *J* = 3.3, 6.0, 10.9 Hz, 1H), 2.20 (m, 1H), 1.55 (s, 3H), 0.80 (s, 9H), -0.07 (s, 3H), 0.00 (s, 3H).

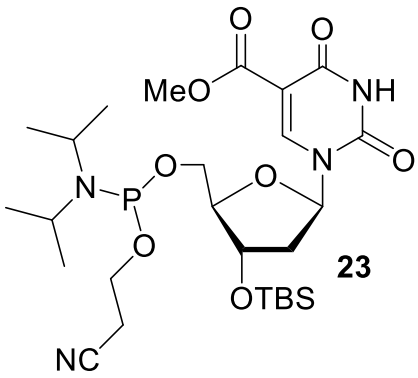
Preparation of 22



22

Compound **21** (811 mg, 1.15 mmol) dissolved in DCM (4.37 mL). A 1 mM stock solution of TCA was prepared by dissolving TCA crystals (462 mg, 2.83 mmol) in H_2O (2.83 mL). The TCA solution (1 M, 1.38 mL) was slowly added, and the reaction turned bright orange. The reaction was monitored by TLC (5% MeOH in DCM, R_f = 0.4, UV active and stained blue with PAA) until the bis-deprotected product began to form after 1 h. The reaction was quenched with sodium bicarbonate solution, diluted with DCM (10 mL), washed with bicarbonate solution (1 x 8 mL), and H_2O (1 x 8 mL). The aqueous layer was extracted with DCM (3 x 10 mL) and the combined organic layers were washed with brine (1 x 30 mL), dried over Na_2SO_4, and concentrated under vacuum. Column chromatography (1.5% MeOH in DCM) was used to purify the crude mixture and resulted in 352 mg (76 %) of **22.** ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.90 (s, 1H), 7.97 (s, 1H), 6.19 (t, *J* = 6.4 Hz, 1H), 4.52 (td, *J* = 3.8, 7.7 Hz, 1H), 4.02 (d, *J* = 6.4 Hz, 1H), 3.98 (d, *J* = 3.8 Hz, 1H), 3.88 (s, 3H), 3.82 (d, *J* = 7.7 Hz, 1H), 2.35 (m, 2H), 0.90 (s, 9H), 0.09 (d, *J* = 0.5 Hz, 6H).

Preparation of 23

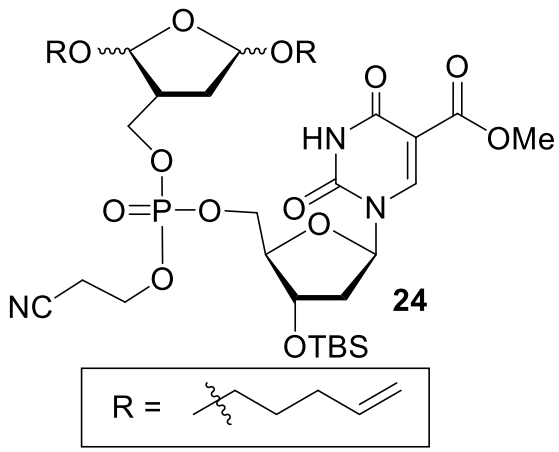


23

Compound **22** (352 mg, 0.875 mmol) was azeotropically dried with pyridine (2 x 2 mL). The flask was cooled to 0 °C. DIPEA (452 mg, 0.64 mL, 3.5 mmol, 4 eq) and DCM (4 mL, 200 mM) were added to dissolve reagent. 2-Cyanoethyl-N,N-diisopropylchlorophosphoramidite (248 mg, 0.234 mL, 1.05 mmol. 1.2 eq) was added and the cold mixture stirred with periodic monitoring by TLC (2:1 Hex/EtOAc, R_f = 0.3, stained with PAA). After 2 h, TLC showed complete conversion to the phosphoramidite. The reaction was diluted with freshly distilled EtOAc (10 mL). The organic layer was washed with saturated bicarbonate solution (2 x15 mL) and the aqueous layers were extracted with distilled EtOAc (2 x 20 mL). The combined organic layers were washed with brine (1 x 25 mL) and dried over Na_2SO_4. The organic layer was concentrated under vacuum and purified by column chromatography (1:1 distilled hexanes/distilled EtOAc), yielding 420 mg (82%) of **23**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.65 (d, *J* = 18.0 Hz, 1H), 6.15 (t, *J* = 6.6 Hz, 1H), 4.80 (s, 1H), 4.50 (m, 1H), 4.16 (d, *J* = 2.0 Hz, 1H), 3.88 (m, 3H), 3.82 (t, *J* = 5.9 Hz, 3H), 3.64 (m, 2H), 2.73 (dt, *J* = 2.0, 5.9 Hz, 2H), 2.39 (m, 1H), 2.30 (dd, *J* = 6.6, 13.3 Hz, 1H), 2.20 (dd, *J* = 6.6, 13.3 Hz, 1H), 1.30 – 1.21 (m, 12H), 0.93 (s, 9H), 0.13 (s, 6H).

^{31}P NMR (400 MHz, $CDCl_3$) δ 148.7, 148.2.

Preparation of 24

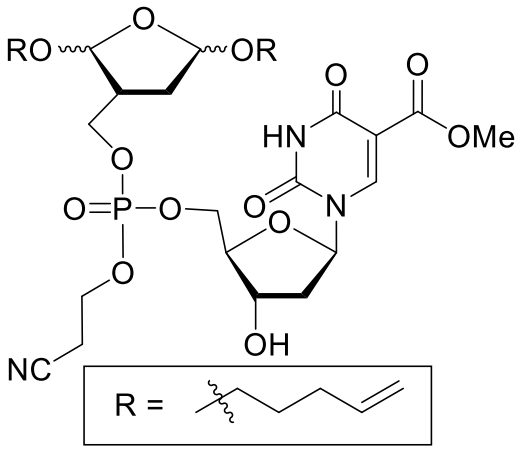


Compounds **9** (234 mg, 0.86 mmol, 1.2 eq) and **23** (420 mg, 0.72 mmol, 1 eq) were azeotropically dried together with toluene (2 x 2 mL). S-Ethyl tetrazole/MeCN (250 mM, 3.44 mL, 0.86 mmol, 1.2 eq) was added to the flask. After 4 h of stirring at 25 ^{o}C, TLC (2:1 EtOAc/Hex, R_f = 0.35, UV active and stained with PAA) indicated consumption of **9**. tBuOOH (5 M, 0.4 mL, 2.16 mmol, 3 eq) was added for 15-20 min. The contents were concentrated under vacuum and column chromatography (2:1 EtOAc/Hex) yielded 460 mg of **24** (70%). ^{1}H NMR (300 MHz, CD_3OD) δ 8.54 (s, 1H), 6.19 (d, *J* = 6.6 Hz, 1H), 5.80 (m, 2H), 5.00 – 4.90 (m, 6H), 4.62 (s, 2H), 4.56 (d, *J* = 13.4 Hz, 1H), 4.34 (m, 4H), 4.13 (m, 2H), 3.82 (s, 3H), 3.68 (m, 2H), 3.40 (m, 2H), 2.90 (m, 2H), 2.37 (m, 2H), 2.19 – 2.05 (m, 4H), 1.91 (m, 1H), 1.64 (dd, *J* = 6.6, 13.4 Hz, 4H), 0.93 (s, 9H), 0.14 (s, 3H), 0.13 (s, 3H).

^{13}C NMR (400 MHz, CD_3OD) δ 172.1, 150.4, 138.9, 115.0, 105.1, 66.3, 60.9, 52.2, 49.3, 49.1, 48.89, 48.88, 48.87, 48.68, 48.67, 48.66, 48.5, 48.4, 48.2, 48.0, 32.2, 31.0, 30.9, 29.6, 25.9, 23.2, 20.6, 19.8, 18.3, 15.2, 14.20, 14.16, -4.8, -5.0.

^{31}P NMR (400 MHz, CD_3OD) δ 8.25, -1.75, -1.82, -1.84, -1.86, -1.90, -1.92, -1.98.

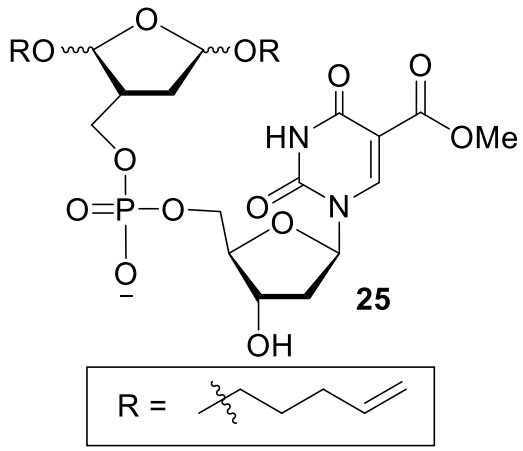
Preparation of 25



Compound **24** (400.8 mg, 0.5 mmol) was dissolved in THF (5 mL, 100 mM). TEA•3HF (400 µL, 2.5 mmol, 5 eq) was added. After confirmed complete by TLC (10% MeOH in DCM, R_f = 0.4, UV active and stained with PAA), the reaction was diluted with DCM (5 mL) and washed with H_2O (1 x 5 mL), sat. sodium bicarbonate (1 x 5 mL), and brine (1 x 5 mL). The combined organic layers dried over Na_2SO_4 and concentrated under vacuum. Column chromatography (5% MeOH in DCM) yielded 297.4 mg of the desilylated intermediate (87%). ^{1}H NMR (400 MHz, CD_3OD) δ 8.52 (s, 1H), 6.16 (t, *J* = 6.1 Hz, 1H), 5.87 – 5.71 (m, 2H), 5.25 – 4.84 (m, 6H), 4.68 (s, 3H), 4.40 (t, *J* = 2.8 Hz, 1H), 4.28 (dd, *J* = 2.8, 6.1 Hz, 4H), 4.12 (s, 2H), 3.78 (s, 3H), 3.63 (m, 2H), 3.33 (m, 2H), 2.86 (m, 2H), 2.41 (m, 1H), 2.24 (m, 1H), 2.05 (m, 4H), 1.81 (m, 1H), 1.60 (m, 4H).

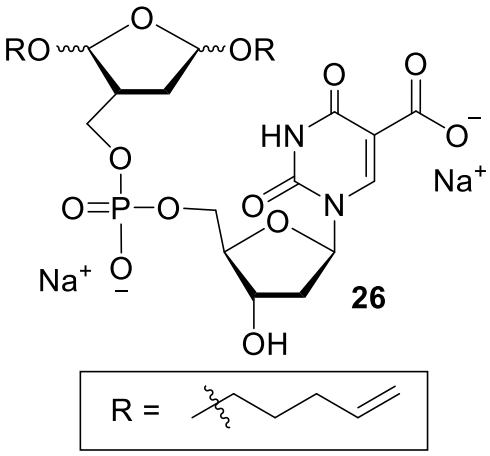
^{13}C NMR (400 MHz, CD_3OD) δ 140.0, 138.9, 117.3, 114.9, 114.8, 105.8, 105.1, 82.4, 68.2, 57.9, 54.4, 54.0, 52.2, 50.3, 49.5, 49.4, 49.3, 49.2, 49.1, 48.9, 48.7, 48.4, 48.2, 48.1, 48.0, 42.9, 30.9, 30.8, 30.7, 29.5, 28.5, 19.8.

^{31}P NMR (400 MHz, CD_3OD) δ -1.90, -1.91, -1.94, -1.95.



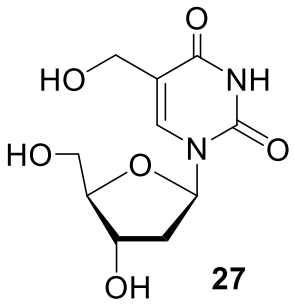
The desilylated intermediate (50 mg, 0.075 mmol) was dissolved in DMF (187.5 µL) and DIPEA (187.5 µL, 50% final solution). The reaction flask was heated at 50 ºC for 25 h until complete as seen by TLC (15% MeOH in DCM, R_f = 0.1, UV active and stained with PAA). The reaction was cooled and concentrated under vacuum. Column chromatography (15% MeOH in DCM) provided 41.2 mg of **25** (89%). ^{1}H NMR (400 MHz, CD_3OD) δ 8.56 (s, 1H), 6.04 (s, 1H), 5.80 (m, 2H), 5.10 – 4.85 (m, 6H), 4.39 (s, 1H), 4.01 (s, 3H), 3.74 (s, 3H), 3.55 (m, 4H), 3.34 (m, 2H), 2.99 (s, 1H), 2.29 (m, 2H), 2.20 (s, 1H), 2.05 (s, 4H), 1.84 (s, 1H), 1.58 (s, 4H).

Preparation of 26



Compound **25** (27.6 mg, 44.7 mmol) and K_2CO_3 (9.6 mg, 67.1 mmol, 1.5 eq) were dissolved in a 1:1 mixture of MeOH/H_2O (300 µL, 150 mM). The reaction stirred for 20 h until complete by TLC (15% MeOH in DCM, 1% TEA, R_f = 0.35, UV active and stained by PAA). The crude mixture passed through a DOWEX Na^+ ion exchange column to yield 23 mg of **26** (86%). ^{1}H NMR (400 MHz, DMSO-*d*$_6$) δ 8.39 (s, 1H), 6.02 (t, *J* = 6.5 Hz, 1H), 5.79 (m, 2H), 5.49 (s, 1H), 5.01 – 4.91 (m, 6H), 4.20 (s, 1H), 3.95 (s, 1H), 3.78 (m, 3H), 3.56 (m, 4H), 3.32 (m, 2H), 2.20 (m, 2H), 2.01 (m, *J* = 6.5 Hz, 4H), 1.78 (s, 1H), 1.55 (d, *J* = 6.5 Hz, 4H).

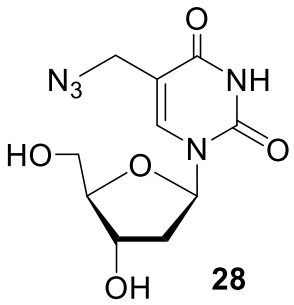
Preparation of 27[156]



27

Deoxyuridine (2.44g, 10.7 mmol) was dissolved in H_2O (15.4 mL). Formaldehyde (5.8 mL, 0.18 mol, 16.7 eq) and KOH (1.15 g, 21.4 mmol, 2 eq) were added to the reaction mixture. The reaction was heated at 65 ^oC and stirred for six days. Over four days, aliquots of KOH solution (4 x 1.15 g, 20 mmol, 2 eq) in H_2O (4 x 5 mL) were added to the reaction to maintain pH 11. Within two days, the reaction turned from colorless to dark brown. After a final 6 days, TLC (20% MeOH in DCM, R_f = 0.35, UV active and stained by PAA) confirmed the conversion of starting material. The reaction was neutralized with DOWEX (500 x 8) H^+ resin. The resin was filtered, washed with H_2O, and the filtrate was concentrated. The crude material was dry loaded onto a silica gel column. Chromatography (10% MeOH, 5% MeCN in DCM) yielded 448 mg (16%) of **27**. ^{1}H NMR (400 MHz, CD_3OD) δ 7.91 (s, 1H), 6.24 (t, *J* = 6.6 Hz, 1H), 4.36 (td, *J* = 3.4, 6.6 Hz, 1H), 4.27 (s, 2H), 3.88 (dd, *J* = 3.4, 6.6 Hz, 1H), 3.74 (dd, *J* = 4.0, 12.0 Hz, 1H), 3.68 (dd, *J* = 4.0, 12.0 Hz, 1H), 2.22 (m, 2H).

Preparation of 28[156]

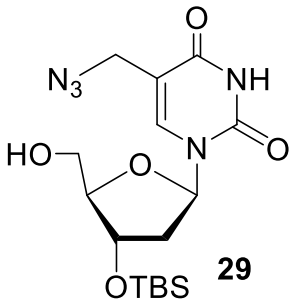


28

Compound **27** (130 mg, 0.51 mmol) was azeotropically dried with pyridine (2 x 2 mL). The starting material liquified once taken off vacuum and stored under Argon. The starting material was dissolved in freshly distilled dioxane (3.4 mL, 150 mM) and TMSCl (0.325 mL, 2.55 mmol, 5 eq) was added. The reaction was heated at 50 ^oC with a water condenser and stirred for 2 h. After 2.5 h, the starting material appeared consumed by TLC (15% MeOH, 5% MeCN in DCM, R_f = 0.45, UV active and stained with PAA) so the reaction mixture was concentrated. The resulting crude was redissolved in DMF (1.4 mL, 350 mM) and NaN_3 (112 mg,

1.53 mmol, 3 eq) was added. After 1 h, TLC (5% MeOH in DCM, R_f = 0.35, UV active and stained with PAA) confirmed conversion. The reaction was concentrated and purified by column chromatography (5% MeOH in DCM) to yield 53 mg (37%) of **28**. ^{1}H NMR (400 MHz, CD_3OD) δ 8.09 (s, 1H), 6.21 (t, *J* = 6.6 Hz, 1H), 4.36 (td, *J* = 3.4, 6.6 Hz, 1H), 4.08 (s, 2H), 3.89 (dd, *J* = 3.4, 6.6 Hz, 1H), 3.76 (dd, *J* = 3.6, 12.1 Hz, 1H), 3.69 (dd, *J* = 3.6, 12.1 Hz, 1H), 2.30 (m, 1H), 2.21 (m, 1H).

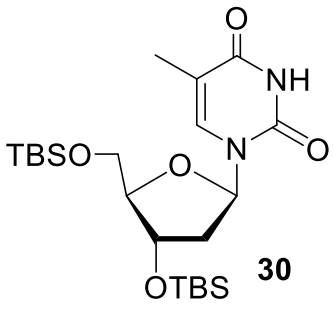
Preparation of 29[189]



29

Compound **31** (38.1 mg, 0.074 mmol, 1 eq, 100 mM) was dissolved in THF (0.8 mL) and cooled to 0 °C. Aqueous TCA (102 mg, 0.625 mmol, 8.5 eq) was prepared in distilled H_2O (0.5 mL) and added to the reaction. After 2.5 h, TLC (2:1 Hex/EtOAc, R_f = 0.3, UV active and stained with PAA) showed conversion to the mono-protected nucleoside and a small amount of the fully deprotected nucleoside. The reaction was quenched with $NaHCO_3$ solid followed by sat. $NaHCO_3$ solution until the pH reached 8. The mixture was diluted with DCM (10 mL) and washed with brine (2 x 10 mL). The aqueous layer was extracted with DCM (2 x 15 mL), the combined organic layers were dried over Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (3% MeOH in DCM) resulting in 13.6 mg (46%) of a clear oil (**29**). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.81 (s, 1H), 6.17 (m, 1H), 4.50 (m, 1H), 4.11 (m, 2H), 3.99 (m, 2H), 3.78 (dd, *J* = 3.3, 6.5 Hz, 1H), 2.28 (dd, *J* = 3.3, 6.5 Hz, 2H), 0.88 (s, 9H), 0.07 (d, *J* = 0.8 Hz, 6H).

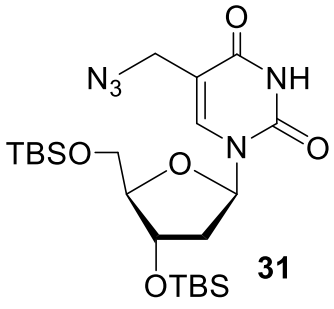
Preparation of 30[190]



30

Thymidine (506 mg, 2.1 mmol, 1 eq) was azeotropically dried with pyridine (3 x 2 mL). Reagent was dissolved in distilled DMF (14 mL, 150 mM). Imidazole (355 mg, 2.5 eq) and TBDMSCl (707 mg, 2.2 eq) were added to the mixture. The reaction was heated under Argon at 50 ºC overnight. When confirmed complete by TLC (5% MeOH in DCM, R_f = 0.6, UV active and stained with PAA), the reaction was cooled and diluted with EtOAc (20 mL). The organic layer was washed with sat. NH_4Cl (3 x 20 mL) and brine (2 x 30 mL). The combined aqueous layers were extracted using EtOAc (2 x 50 mL). The combined organic layers were dried over Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (2:1 hexanes/EtOAc) to yield 720 mg (80%) of **30**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.34 (s, 1H), 7.45 (d, *J* = 1.3 Hz, 1H), 6.31 (m, 1H), 4.38 (m, 1H), 3.91 (d, *J* = 1.3 Hz, 1H), 3.85 (dd, *J* = 2.6, 11.4 Hz, 1H), 3.74 (dd, *J* = 2.6, 11.4 Hz, 1H), 2.23 (dd, *J* = 2.6, 14.0 Hz, 1H), 2.01 (m, 1H), 1.89 (s, 3H), 0.89 (d, *J* = 14.0 Hz, 18H), 0.11 (m, 12H).

Preparation of 31[189]

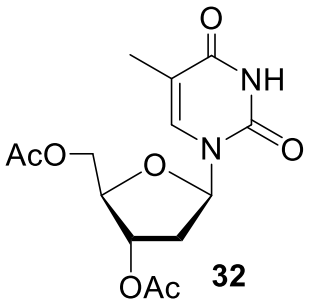


31

Compound **30** (155 mg, 0.33 mmol, 1 eq, 165 mM) was mixed with NBS (105.6 mg, 0.6 mmol, 1.8 eq) and AIBN (10.7 mg, 0.065 mmol, 0.18 eq). Distilled benzene (2 mL) was added to dissolve the reactants. The mixture was sparged with Ar for 20 minutes. The reaction was refluxed (80 ºC) and stirred for 30-45 min, while the mixture turned orange-yellow. When TLC (2:1 Hex/EtOAc, 2% MeOH, R_f = 0.4, UV active and stained with PAA), showed the starting material was mostly converted without the formation of side products, the mixture was immediately filtered through a glass frit and concentrated under vacuum to give a yellow crude residue. The residue was dissolved in DMF (1.5 mL, 200 mM) and stirred under Argon. NaN_3 (44 mg, 0.677

mmol, 2 eq) was added to the solution, turning red in color. The reaction was heated to 60 ºC and stirred for 1-2 h. When TLC (1:1 Hex/EtOAc, 2% MeOH, UV active and stained with PAA), showed the starting material was mostly converted without the formation of side products, the reaction was cooled and diluted with EtOAc (30 mL), washed with water (2 x 20 mL), and brine (2 x 20 mL). The aqueous layer was extracted with EtOAc (2 x 40 mL), and the combined organic layers were dried over Na_2SO_4 and concentrated under vacuum to give a red crude residue. The residue was purified by column chromatography (7:1 DCM/EtOAc) to give 65.9 mg (39.2%) of **31**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.71 (s, 1H), 7.46 (d, *J* = 2.3 Hz, 1H), 6.30 (dd, *J* = 5.8, 13.6 Hz, 1H), 4.39 (dd, *J* = 2.3, 3.5 Hz, 1H), 4.21 – 4.01 (m, 1H), 3.95 (m, 2H), 3.86 (td, *J* = 3.5, 11.4 Hz, 1H), 3.76 (td, *J* = 2.3, 11.4 Hz, 1H), 2.27 (ddd, *J* = 2.3, 5.8, 13.6 Hz, 1H), 2.03 – 1.95 (m, 1H), 0.96 – 0.84 (m, 18H), 0.24 – 0.02 (m, 12H).

Preparation of 32[191]

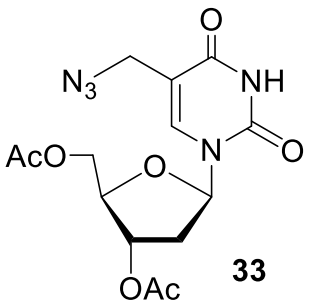


32

Thymidine (3.08 g, 12.7 mmol, 1 eq) was azeotropically dried with pyridine (2 x 8 mL). DMAP (159 mg, 1.27 mmol, 0.2 eq) was added to a flask and flushed with Argon. The reagents were dissolved in distilled DCM (6.7 mL) and cooled to 0 ºC. Pyridine (12.5 g, 12.8 mL, 158.8 mmol, 12.5 eq) was added slowly to the solution. Next, acetic anhydride (3.82 g, 6 mL, 63.6 mmol, 5 eq) was added dropwise. After 3.5 h, TLC (3:2 EtOAc/Hex, R_f = 0.8, stained with PAA) confirmed the reaction complete, the reaction was quenched with sat. $NaHCO_3$ until neutral pH was reached. The mixture was diluted with EtOAc (15 mL) and washed with water (2 x 15 mL). The aqueous layer was extracted with EtOAc (2 x 30 mL) and the combined organic layers were washed with brine (1 x 50 mL) and dried over Na_2SO_4. The crude residue was purified by column chromatography (1:1 Hex/EtOAc) to yield 3.8 g (93%) of **32.** ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.10 (s, 1H), 7.21 (s, 1H), 6.24 (dd, *J* = 4.2, 8.4 Hz, 1H), 5.13 (d, *J* = 6.8 Hz, 1H), 4.28

(dd, *J* = 4.2, 12.1 Hz, 1H), 4.23 (dd, *J* = 3.3, 12.1 Hz, 1H), 4.15 (dd, *J* = 3.3, 6.8 Hz, 1H), 2.37 (dd, *J* = 6.0, 8.4 Hz, 1H), 2.13 (m, 1H), 2.03 (d, *J* = 6.0 Hz, 6H), 1.83 (s, 3H).

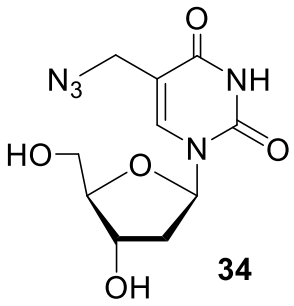
Preparation of 33[192]



33

Compound **32** (785.8 mg, 2.44 mmol, 1 eq) was mixed with NBS (693.9 mg, 3.9 mmol, 1.8 eq) and AIBN (64 mg, 0.39 mmol, 0.18 eq). Distilled benzene (16 mL) dissolved the reactants. The mixture was sparged with Ar for 20 min. The reaction was heated to reflux (80 °C) and stirred for 30-45 min while the mixture turned red. When TLC (2:1 EtOAc/Hex, R_f = 0.6, UV active and stained with PAA), showed the starting material was mostly consumed without the formation of side products, the mixture was immediately filtered through a glass frit and concentrated under vacuum to give a yellow crude residue. The residue was redissolved in DMF (12 mL, 200 mM). NaN_3 (904 mg, 4.88 mmol, 2 eq) was added to the solution, which turned brown. The reaction was heated to 60 °C and stirred for 1-2 h. When TLC (2:1 EtOAc/Hex, R_f = 0.2, UV active and stained with PAA), showed the disappearance of the intermediate, the reaction was cooled and diluted with EtOAc (30 mL), washed with water (2 x 40 mL), and brine (2 x 40 mL). The aqueous layer was extracted with EtOAc (2 x 80 mL) and the combined organic layers were dried over Na_2SO_4 and concentrated under vacuum to give a red crude residue. The residue was purified by column chromatography in 1:1 Hex/EtOAc to give 628 mg (55%) of **33**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.57 (s, 1H), 6.29 (dd, *J* = 5.6, 8.5 Hz, 1H), 5.22 (dt, *J* = 1.9, 8.5 Hz, 1H), 4.40 (dd, *J* = 4.2, 12.1 Hz, 1H), 4.31 (dd, *J* = 4.2, 12.1 Hz, 1H), 4.28 – 4.26 (m, 1H), 4.22 (dd, *J* = 1.0, 14.4 Hz, 1H), 4.12 (dd, *J* = 1.0, 14.4 Hz, 1H), 2.52 (ddd, *J* = 1.9, 5.6, 14.4 Hz, 1H), 2.30 – 2.18 (m, 1H), 2.12 (d, *J* = 5.6 Hz, 6H).

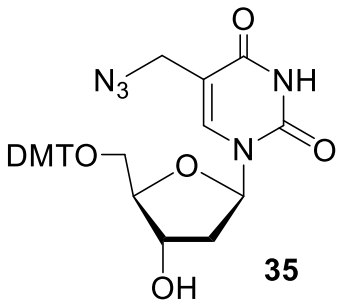
Preparation of 34[193]



34

Compound **33** (300 mg, 0.82 mmol) was dissolved in dioxane (12 mL, 68 mM). Concentrated NH_4OH (29% in water, 12 mL, 18 M) was added to the reaction ([**33**] = 34 mM, [NH_4OH] = 9 M). The reaction was stirred for 18 h. Analysis by TLC (4:1 EtOAc/Hex, R_f = 0.05, UV active and stained with PAA) confirmed the absence of starting material. The reaction was concentrated and purified by column chromatography (4:1 EtOAc/Hex -> 100 % EtOAc, dry loaded) to yield 127 mg (50%) of **34**. ^{1}H NMR (400 MHz, CD_3OD) δ 8.09 (s, 1H), 6.21 (t, *J* = 6.6 Hz, 1H), 4.36 (td, *J* = 3.4, 6.6 Hz, 1H), 4.03 (s, 2H) 3.89 (dd, *J* = 3.4, 6.2 Hz, 1H), 3.76 (dd, *J* = 3.4, 12.1 Hz, 1H), 3.69 (dd, *J* = 3.7, 12.1 Hz, 1H), 2.26 (dd, *J* = 3.7, 6.2 Hz, 1H), 2.19 (m, 1H).

Preparation of 35[194]

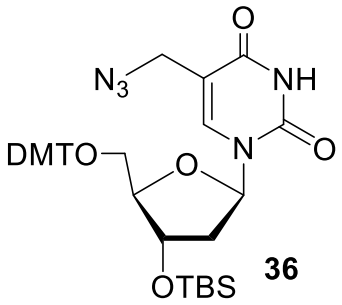


35

Compound **34** (895.5 mg, 3.16 mmol) was azeotropically dried with pyridine (2 x 3 mL). DMT-Cl (1.32 g, 3.79 mmol, 1.2 eq) was added and reactants were dissolved in pyridine (6.4 mL, 500 mM) at 0 °C. The mixture turned orange and was monitored by TLC (5:1 EtOAc/Hex, R_f = 0.3, UV active and stained orange then blue with PAA). After 2 h, a side product (presumably the bis-tritylated product) was observed via TLC. Therefore, the reaction was concentrated under vacuum. The residue was dissolved in EtOAc (30 mL) and washed with sat. sodium bicarbonate (1 x 30 mL) and H_2O (1 x 30 mL). The aqueous layer was extracted with EtOAc (2 x 30 mL) and the combined organic layers were washed with brine (1 x 50 mL), dried over Na_2SO_4, and concentrated under vacuum. The crude material was purified by column chromatography (5:1 EtOAc/hexanes) to yield 1.078 g (60%) of **35.** ^{1}H NMR

(400 MHz, $CDCl_3$) δ 8.56 (s, 1H), 7.82 (s, 1H), 7.65 (m, 1H), 7.36 (m, 1H), 7.30 – 7.22 (m, 5H), 6.80 (m, 4H), 6.37 (d, *J* = 5.8 Hz, 1H), 4.57 (t, *J* = 3.0 Hz, 1H), 4.06 (s, 2H), 3.76 (s, 6H), 3.52 (d, *J* = 13.7 Hz, 1H), 3.45 (d, *J* = 3.2 Hz, 1H), 3.40 – 3.34 (dd, *J* = 3.2, 13.7 Hz, 2H), 3.32 (d, *J* = 13.5 Hz, 2H), 2.44 (ddd, *J* = 3.0, 5.8, 13.5 Hz, 1H), 2.34 – 2.24 (m, 1H).

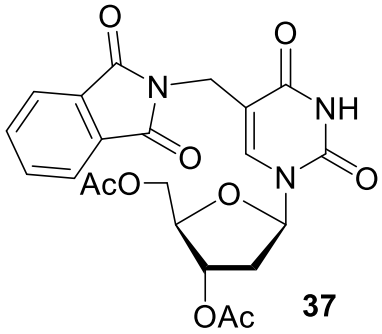
Preparation of 36[194]



36

Compound **35** (452.2 mg, 0.77 mmol, 1 eq) was azeotropically dried with pyridine (3 x 1 mL). Distilled DMF (2.6 mL, 300 mM) dissolved **35.** Imidazole (330. mg, 4.64 mmol, 6 eq) and TBDMSCl (377 mg, 2.32, 3 eq) were added to the mixture. The reaction was heated at 50° C and stirred overnight. When confirmed complete by TLC (2:1 Hex/EtOAc, R_f = 0.25, UV active and stained with PAA), the reaction was cooled and diluted with EtOAc (20 mL). The organic layer was washed with sat. NH_4Cl (3 x 20 mL) and brine (2 x 30 mL). The combined aqueous layers were extracted using EtOAc (2 x 50 mL). The combined organic layers dried over Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (2:1 hexanes/EtOAc) yielding 289 mg (55%) of **36**. ^1H NMR (400 MHz, $CDCl_3$) δ 7.92 (s, 1H), 7.35 (d, *J* = 8 Hz, 2H), 7.24 (dd, *J* = 3.7, 8 Hz, 7H), 6.78 (d, *J* = 8 Hz, 4H), 6.27 (t, *J* = 6.2 Hz, 1H), 4.49 (td, *J* = 2.4, 3.7 Hz, 1H), 4.05 (d, *J* = 3.7 Hz, 1H), 3.75 (s, 6H), 3.56 (d, *J* = 13.7 Hz, 1H), 3.52 (dd, *J* = 2.7, 10.7 Hz, 1H), 3.42 (d, *J* = 13.7 Hz, 1H), 3.33 (dd, *J* = 2.7, 10.7 Hz, 1H), 2.39 (dd, *J* = 3.7, 6.2 Hz, 1H), 2.28 (dd, *J* = 6.2, 13.7 Hz, 1H), 0.85 (s, 9H), 0.07 (s, 3H), 0.02 (s, 6H).

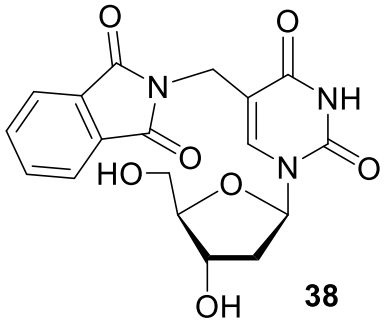
Preparation of 37



37

Compound **32** (1.25 g, 3.8 mmol, 1 eq) was mixed with NBS (1.26 g, 6.9 mmol, 1.8 eq) and AIBN (122 mg, 0.69 mmol, 0.18 eq). Distilled benzene (19 mL, 200 mM) dissolved the reactants. The mixture was sparged with Ar for 20 min. The reaction was heated to reflux (80 °C) and stirred for 30-45 min, during which the mixture turned red. At that time, TLC (4:1 EtOAc/Hex, $R_f = 0.5$, UV active and stained with PAA), showed the starting material was mostly converted without the formation of side products, the mixture was immediately filtered through a glass frit and concentrated under vacuum to give a yellow crude residue. The residue was redissolved in DMF (15 mL, 200 mM) and stirred under Argon. Potassium phthalimide (1.41 g, 7.6 mmol, 2 eq) was added to the solution, which turned brown in color. The reaction was heated at 60° C. After 2 h, TLC (4:1 EtOAc/Hex, $R_f = 0.3$, UV active and stained with PAA), showed the disappearance of the intermediate. The reaction was cooled, diluted with EtOAc (30 mL), and washed with water (2 x 40 mL) and brine (2 x 40 mL). The aqueous layer was extracted with EtOAc (2 x 70 mL), and the combined organic layers were dried over Na_2SO_4 and concentrated under vacuum to give a red crude residue. The residue was purified by column chromatography (4:1 EtOAc/Hex) to give 523 mg (30%) of **37**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.98 (s, 1H), 7.85 – 7.61 (m, 5H), 6.24 (dd, J = 5.9, 8.6 Hz, 1H), 5.20 (d, J = 5.9 Hz, 1H), 4.57 (m, 2H), 4.41 – 4.28 (m, 1H), 4.28 – 4.15 (m, 2H), 2.52 – 2.40 (m, 1H), 2.24 – 2.12 (m, 1H), 2.01 (d, J = 1.3 Hz, 6H).

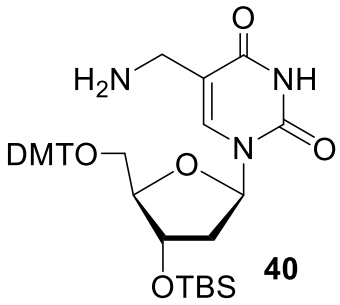
Preparation of 38



38

Compound **37** (430 mg, 0.91 mmol) was dissolved in MeOH (3.67 mL) and triethylamine (648 mg, 0.89 mL, 6.4 mmol, 7 eq). The reaction was stirred at 25 °C for 26 h until complete by TLC (4:1 EtOAc/Hex, R_f = 0.1, UV active and stained with PAA). The crude material was purified by column chromatography (4:1 EtOAc/Hex -> 100% EtOAc) to give 152 mg (45%) of **38**. ^{1}H NMR (400 MHz, CD_3OD) δ 8.00 (s, 1H), 7.83 (dd, *J* = 3.0, 5.6 Hz, 2H), 7.72 (dd, *J* = 3.0, 5.6 Hz, 2H), 6.24 (t, *J* = 6.4 Hz, 1H), 4.55 (m, 2H), 4.00 (s, 1H), 3.96 (s, 1H), 3.78 (s, 1H), 3.64 (1H), 2.38 (s, 1H), 2.31 (d, *J* = 6.4 Hz, 1H).

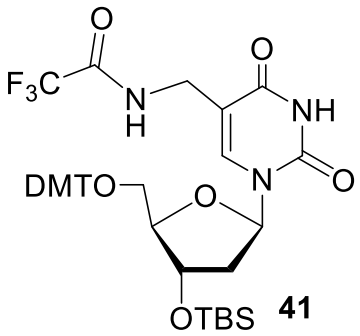
Preparation of 40



40

Compound **36** (153.9 mg, 0.261 mmol) dissolved in a solution of 60% MeOH, 20% tBuOH, 20% H_2O (5.2 mL, 50 mM). Activated palladium on carbon (102.4 mg, 60 wt%) was added and the mixture was flushed with a Hydrogen balloon three times. The reaction was sparged with H_2 for several hours, with multiple rounds of refilling the balloon. After 4 h, TLC (1:1 Hex/EtOAc, UV active and stained with PAA) confirmed the starting material (R_f= 0.6) converted to the product (R_f= 0.02). The crude mixture was filtered through Celite and concentrated to yield 139.7 mg (93%) of **23**. ^{1}H NMR (400 MHz, CD_3OD) δ 7.85 (s, 1H), 7.42 (d, *J* = 6.8 Hz, 2H), 7.36 – 7.21 (m, 7H), 6.97 – 6.71 (m, 4H), 6.24 (m, 1H), 4.52 (t, *J* = 4.0 Hz, 1H), 3.94 (dd, *J* = 4.0, 6.8 Hz, 1H), 3.82 (dd, *J* = 3.2, 10.7 Hz, 1H), 3.76 (s, 6H), 3.69 (d, *J* = 4.0 Hz, 1H), 3.49 (dd, *J* = 3.2, 10.7 Hz, 1H), 3.18 – 3.03 (m, 1H), 2.40 – 2.24 (m, 2H), 0.82 (s, 9H), 0.00 (s, 3H), -0.03 (s, 3H).

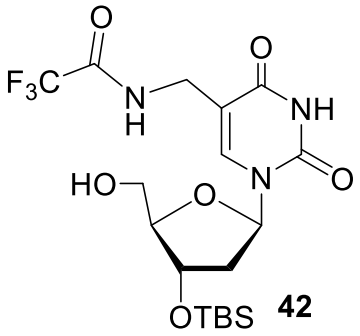
Preparation of 41



41

Crude **40** (83.8 mg, 0.12 mmol) was dissolved in freshly distilled THF (2.48 mL, 50 mM). Distilled triethylamine (0.26 mL, 1.86 mmol, 15 eq) and ethyl trifluoroacetate (0.15 mL, 1.24 mmol, 10 eq) was added to the flask. The reaction was stirred for 4 hours at 25 °C. TLC (1:1 Hex/EtOAc, R_f = 0.5, UV active and stained orange then blue by PAA) confirmed conversion to product. The reaction concentrated to yield 65.4 mg (70%) of **41**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.82 (s, 1H), 7.40 (d, *J* = 7.5 Hz, 2H), 7.35 – 7.09 (m, 7H), 6.81 (d, *J* = 7.5 Hz, 4H), 6.24 (t, *J* = 6.4 Hz, 1H), 4.40 (t, *J* = 2.8 Hz, 1H), 3.93 (d, *J* = 3.4 Hz, 1H), 3.74 (s, 6H), 3.69 (t, *J* = 6.4 Hz, 3H), 3.37 (dd, *J* = 2.8, 10.6 Hz, 1H), 3.29 (dd, *J* = 3.4, 10.6 Hz, 1H), 2.28 (m, 1H), 2.14 (m, 1H), 0.79 (s, 9H), 0.01 (s, 3H), -0.04 (s, 3H).

Preparation of 42



42

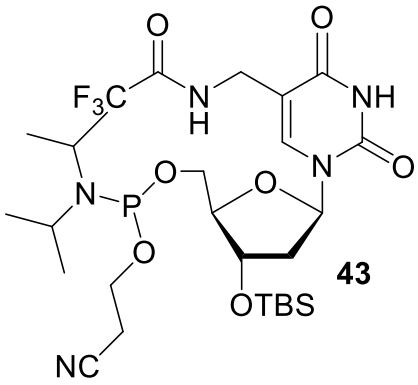
Crude **41** (116.9 mg, 0.15 mmol) was dissolved in 3% TCA in DCM (2.48 mL, 50 mM). The reaction stirred for 1.5 h until TLC (1:1 hex/EtOAc, R_f = 0.2, UV active and stained blue with PAA) showed starting material conversion to product. The reaction was quenched with sat. sodium bicarbonate (3 mL) and diluted with DCM (5 mL). The organic layer was washed with bicarbonate solution (1 x 8 mL). The aqueous layer was extracted with DCM (3 x 10 mL) and the combined organic layers were washed with brine (1 x 40 mL), dried with Na_2SO_4, and concentrated. The crude mixture was purified by column chromatography (2:1 Hex/EtOAc) to yield 59 mg (64%) of **42**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.44 (s, 1H), 8.26 (s, 1H), 6.20 (t, *J* = 6.4 Hz, 1H), 4.47 (q, *J* = 3.2 Hz, 1H), 4.17 (d, *J* = 6.4 Hz, 2H),

3.97 (s, 1H), 3.96 (dd, *J* = 3.2, 9.8 Hz, 2H), 3.76 (d, *J* = 9.8 Hz, 1H), 2.42 – 2.29 (m, 1H), 2.25 – 2.11 (m, 1H), 0.88 (m, 9H), 0.11 – 0.00 (m, 6H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 164.3, 158.5, 158.1, 150.4, 141.8, 117.2, 114.3, 109.1, 88.5, 86.7, 77.4, 77.1, 76.8, 71.8, 61.8, 41.9, 36.6, 25.7, 21.1, 18.0, -4.78, -4.9.

ESI-TOF m/z calculated for $C_{18}H_{28}F_3N_3O_6Si$ (M + H) – 468.1699, 468.1766 observed.

Preparation of 43



43

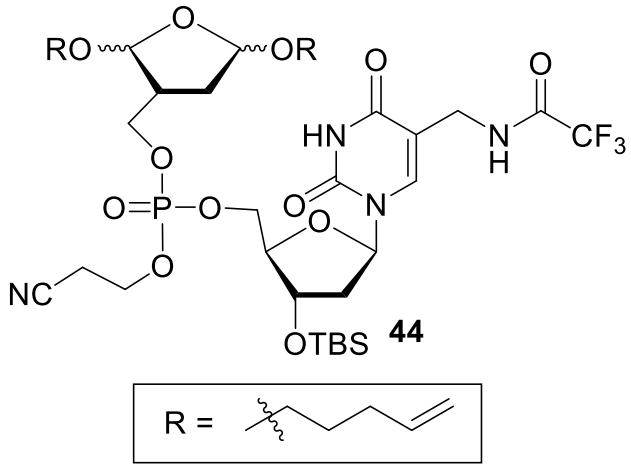
Compound **42** (59 mg, 0.126) was azeotropically dried with distilled pyridine (2 x 0.5 mL). The flask was cooled to 0 ºC. DIPEA (65 mg, 90 µL, 0.504 mmol, 4 eq) was added to the cold starting material and the reactants dissolved in DCM (1.25 mL, 100 mM). 2-Cyanoethyl-N, N-diisopropylchlorophosphor-amidite (35.7 mg, 34 µL, 0.151 mmol. 1.2 eq) was added and the reaction was monitored periodically by TLC (1:1 Hex/EtOAc, R_f = 0.6, stained with PAA). After 2 h, TLC showed complete conversion to the phosphoramidite. The reaction was diluted with freshly distilled EtOAc (3 mL). The organic layer was washed with sat. bicarbonate solution (2 x 5 mL) and the combined aqueous layers were extracted with distilled EtOAc (2 x 10 mL). The combined organic layers were washed with brine (1 x 25 mL) and dried over Na_2SO_4. The organic layers were concentrated under vacuum and the crude material was purified by column chromatography (3:1 distilled hexanes/distilled EtOAc) yielding 44.3 mg (55%) of **43**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.02 (s, 1H), 6.31 (t, *J* = 6.5 Hz, 1H), 4.44 (dd, *J* = 3.4, 6.6 Hz, 1H), 4.19 (dd, *J* = 3.4, 13.3 Hz, 1H), 4.03 (d, *J* = 2.4 Hz, 2H), 3.93 (t, *J* = 2.4 Hz, 2H), 3.82 (m, 2H), 3.59 (m, 2H), 2.67 (t, *J* = 6.6 Hz, 2H), 2.40 – 2.22 (m, 2H), 2.11 (dt, *J* = 6.5, 13.3 Hz, 1H), 1.26 – 1.14 (m, 12H), 0.88 (s, 9H), 0.13 – 0.03 (m, 6H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 171.1, 163.84, 163.81, 150.1, 150.0, 140.1, 117.1, 109.2, 109.1, 85.2, 77.3, 77.0, 76.6, 72.4, 71.7, 60.3, 58.5, 43.1, 43.0, 41.1, 34.6, 31.5, 29.0, 25.6, 25.2, 24.65, 24.57, 24.5, 22.6, 20.96, 20.4, 20.3, 18.7, 17.8, 14.1, 14.0, 11.3.

^{31}P NMR (400 MHz, $CDCl_3$) δ 148.9, 148.7.

ESI-TOF m/z calculated for $C_{27}H_{45}F_3N_5O_7PSi$ (M + Na) − 690.2778, 690.2438 observed.

Preparation of 44



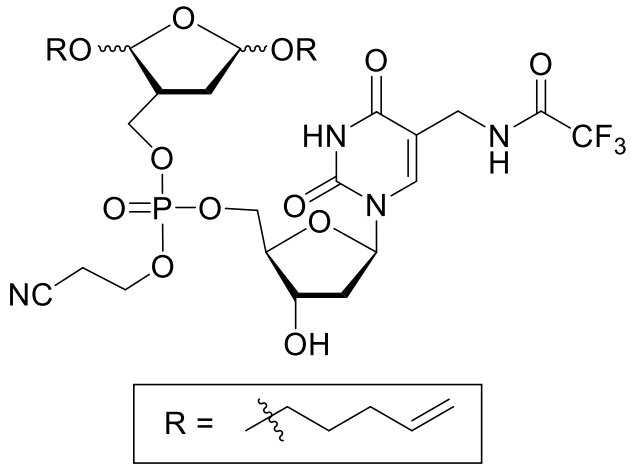
Compounds **43** (44.3 mg, 0.068 mmol, 1 eq) and **9** (29.7 mg, 0.082 mmol, 1.2 eq) were azeotropically dried together with toluene (2 x 2 mL). S-Ethyl tetrazole/MeCN (250 mM, 330 µL, 0.082 mmol, 1.2 eq) was added to the flask. After 3.5 h, TLC (2.5:1 EtOAc/Hex, R_f = 0.6, UV active and stained with PAA) indicated that the majority of **9** was consumed. tBuOOH (500 mM, 40.8 µL, 0.204 mmol, 3 eq) was added for 15-20 min. The reaction was concentrated under vacuum and column chromatography (1:1 EtOAc/Hex -> 3:1 EtOAc/Hex) yielded 43 mg (75 %) of **44**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.61 (m, 1H), 8.15 (s, 1H), 7.76 (s, 1H), 6.28 (q, *J* = 6.8 Hz, 1H), 5.80 (m, 2H), 5.58 (m, 1H), 5.65 – 4.89 (m, 6H), 4.77 (d, *J* = 3.4 Hz, 1H), 4.40 (d, *J* = 3.4 Hz, 1H), 4.36 – 4.25 (m, 5H), 4.02 (s, 1H), 3.68 (dd, *J* = 3.2, 6.0 Hz, 2H), 3.54 (m, 1H), 2.77 (t, *J* = 6.0 Hz, 2H), 2.32 (m, 1H), 2.10 (d, *J* = 6.8 Hz, 5H), 1.84 (m, 1H), 1.73 – 1.61 (m, 6H), 0.87 (s, 9H), 0.08 (d, *J* = 1.9 Hz, 6H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 163.4, 150.1, 138.1, 114.83, 114.8, 110.8, 108.0, 85.1, 68.7, 67.3, 67.1, 67.0, 62.7, 31.6, 30.2, 29.0, 28.9, 25.6, 22.6, 21.1, 17.8, 14.2, 14.1, 11.4, -4.8, -5.0.

^{31}P NMR (400 MHz, $CDCl_3$) δ -1.8, -1.9.

ESI-TOF m/z calculated for $C_{36}H_{56}F_3N_4O_{12}PSi$ (M - H) − 851.3354, 851.3229 observed.

Preparation of 45

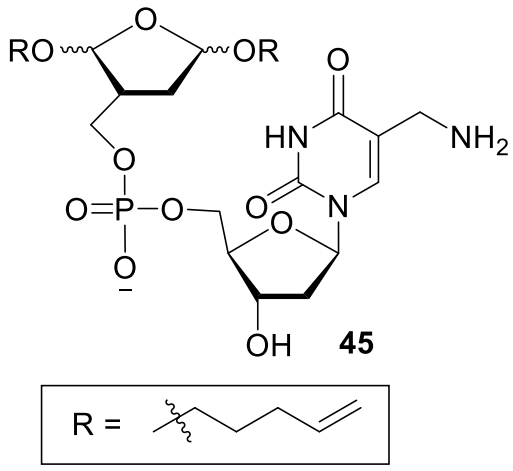


Compound **44** (28.5 mg, 0.033 mmol) was dissolved in THF (330 µL, 100 mM). TEA•3HF (28 µL, 0.17 mmol, 5 eq) was added to the flask and the reaction was stirred at 25 °C overnight. After confirmed complete by TLC (3:1 EtOAc/Hex, R_f = 0.1, UV active and stained with PAA), the reaction was diluted with DCM (2 mL) and washed with H_2O (1 x 3 mL), sat. sodium bicarbonate (1 x 3 mL), and brine (1 x 3 mL). The combined organic layers were dried over Na_2SO_4 and concentrated under vacuum. Column chromatography (4:1 EtOAc/Hex -> 100% EtOAc) provided 22 mg (92 %) of the desilylated intermediate. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.75 (s, 1H), 8.09 (s, 1H), 7.76 (s, 1H), 6.25 (s, 1H), 5.80 (dd, *J* = 6.2, 10.3 Hz, 2H), 5.60 (m. 1H), 5.10-4.90 (m, 6H), 4.77 (d, *J* = 10.3 Hz, 1H), 4.49 (s, 1H), 4.32 (s, 4H), 4.21 (s, 2H), 4.09 (s, 2H), 3.68 (m, 2H), 3.60 – 3.46 (m, 1H), 3.40 (s, 1H), 2.80 (d, *J* = 6.2 Hz, 2H), 2.41 (s, 1H), 2.16 (s, 2H), 2.09 (m, 4H), 1.75 – 1.56 (m, 4H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 137.9, 131.1, 114.7, 107.7, 105.2, 77.3, 77.0, 76.7, 53.4, 50.3, 46.0, 30.0, 28.8, 25.6, 8.4.

^{31}P NMR (400 MHz, $CDCl_3$) δ -1.81, -1.84, -1.9, -2.1.

ESI-TOF m/z calculated for $C_{30}H_{42}F_3N_4O_{12}P$ (M -H) − 737.2489, 737.2353 observed.



The desilylated intermediate (27.5 mg, 0.037 mmol) was dissolved in concentrated aqueous ammonia (560 µL). The reaction was capped and stirred at 0 °C overnight. After 10 h, TLC (15% MeOH in DCM, R_f = 0.2, UV active and stained with $KMnO_4$) confirmed the disappearance of starting material. The reaction was concentrated to yield 22.1 mg (100%) of **45**. The desired product confirmed by ESI-MS m/z calculated for $C_{25}H_{40}N_3O_{11}P$ (M -H)− 588.16, 588.2 observed. ESI-TOF m/z calculated for $C_{25}H_{40}N_3O_{11}P$ (M -H)− 588.2400, 588.2308 observed.

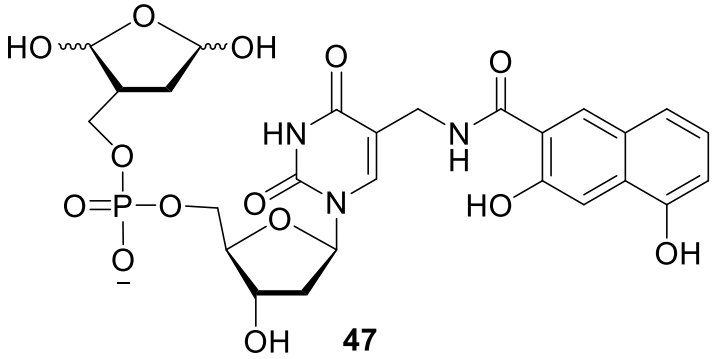
General Procedure for the preparation of Library 45

Amine scaffold **46** (100 nmol) was azeotropically dried with carboxylic acid (140 nmol, 1.4 eq) in pyridine (1 x 15 µL) using a Speed Vac concentrator in a 384-well microtiter plate (VWR). To each well, a 2X activating solution (5 µL; 28 mM HBTU and 28 mM HOBt in DMF), DIPEA (2 µL), and DMF (3 µL) were added. The final concentrations during reaction were: [**46**] = 10 mM, [acid] = 14 mM, [HBTU] = 14 mM, [HOBt] = 14 mM, 20% DIPEA in DMF. The well plate was shaken at 25 °C overnight. Some wells were analyzed by ESI-MS to confirm coupling efficiency. The solutions were concentrated using a Speed Vac concentrator and the well plate was covered and stored at -80 °C.

Immediately before an assay, the amide was thawed, dissolved in DMF (4 µL, 25 mM). An aliquot (2 µL, 50 nmol) was mixed with NBS (8 µL, 15 mM, 2.4 eq, 40% H_2O in MeCN) at 0 °C for 9 min. The concentrations during reaction were: [SM] = 5 mM, [NBS] = 12 mM, 20% DMF, 3% H_2O in MeCN. After 4 min, sat. $Na_2S_2O_3$ (5 µL) was added and reaction quenched on ice for

10 min. Samples were concentrated with a Speed Vac concentrator. Some samples were analyzed ESI-MS to confirm product formation.

Preparation of 47



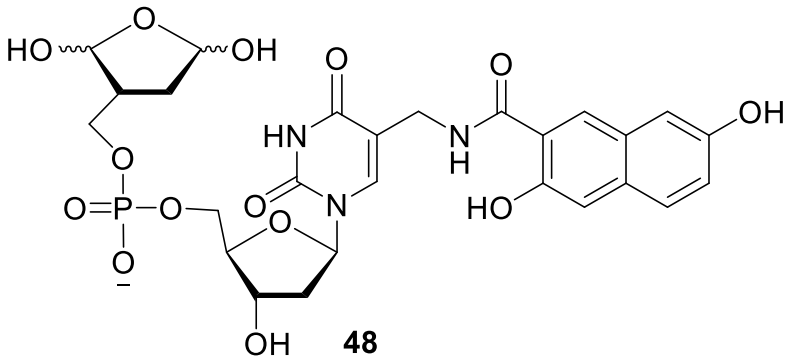
47

A stock solution of **54** (20 mM) in 6% H_2O in MeCN was prepared and stored at -20 °C. An aliquot of the stock solution (5 µL, 100 nmol) was cooled to 0 °C. A solution of recrystallized NBS (5 µL, 30 mM, 300 nmol, 3 eq) was added and the reaction was stirred for 4 min (final: [**54**] = 5 mM, 3% H_2O, [NBS] = 15 mM). After 4 min, saturated $Na_2S_2O_3$ (10 µL) was added and the reaction was stirred for 10 min to quench it. The reaction was concentrated, dissolved in 1:1 MeCN/H_2O, and analyzed by ESI.

ESI-MS m/z calculated for $C_{26}H_{30}N_3O_{14}P$ (M-1) 638.14 calculated, 638.24 observed.

$\varepsilon_{364} = 1.72 \times 10^3\ M^{-1}cm^{-1}$

Preparation of 48



48

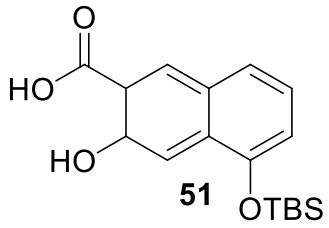
A stock solution of **53** (20 mM) in 6% H_2O in MeCN was prepared and stored at -20 °C. An aliquot of the stock solution (5 µL, 100 nmol) was cooled to 0 °C. A solution of recrystallized NBS (5 µL, 30 mM, 300 nmol, 3 eq) was added and the reaction stirred for 4 min (final: [**53**] = 5 mM, 3% H_2O, [NBS] = 15 mM). After 4 min, saturated $Na_2S_2O_3$ (10 µL)

was added and the reaction was stirred for 10 min to quench it. The reaction was concentrated, dissolved in 1:1 MeCN/H_2O, and analyzed by ESI.

ESI-MS m/z calculated for $C_{26}H_{28}N_3O_{13}P$ (M - H_2O - H)− 621.14 calculated, 621.16 observed.

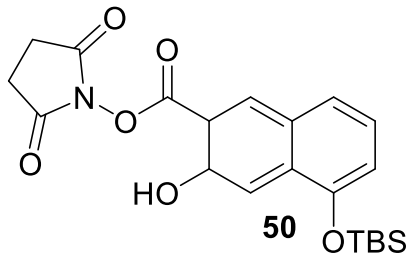
$\varepsilon_{380} = 1.35 \times 10^3\ M^{-1}cm^{-1}$

Preparation of 49[195]



3,5-Dihydroxy naphthoic-2-carboxylic acid (485 mg, 2.37 mmol) was azeotropically dried with pyridine (3 x 1 mL). TBDMSCl (1.09 g, 7.11 mmol, 3 eq) and imidazole (967 mg, 14.2 mmol, 6 eq) were added to the flask. The contents were dissolved in DMF (9.5 mL, 250 mM) and reaction was heated at 60 ºC and stirred overnight. The following morning, TLC (3% MeOH in DCM) suggested starting material was consumed so H_2O (0.2 mL) was added, and the reaction was stirred for 2 h. Afterwards, the reaction was concentrated and purified by column chromatography (4:1 Hex/EtOAc -> 1:1 Hex/EtOAc) to yield 400 mg (53%) of **49**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.46 (s, 1H), 7.44 (s, 1H), 7.26 (d, *J* = 7.6 Hz, 1H), 6.96 (d, *J* = 7.6 Hz, 1H), 6.72 (d, *J* = 7.6 Hz, 1H), 1.00 (s, 9H), 0.19 (s, 6H).

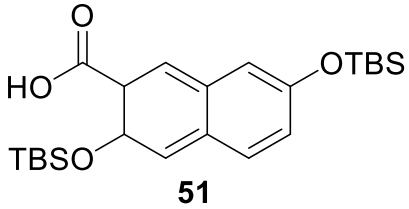
Preparation of 50[195]



Compound **49** (238 mg, 0.75 mmol) was azeotropically dried with pyridine (2 x 1 mL). NHS (128.8 mg, 1.12 mmol, 1.5 eq), and EDC (219 mg, 1.12 mmol, 1.5 eq) were added to the flask and the reagents were dissolved in DMF and stirred overnight at 25 ºC. After 16 h, TLC (5% MeOH in DCM, R_f = 0.7) suggested the SM was completely converted so the reaction was concentrated. The product was purified by column chromatography (3:1 Hex/EtOAc -> 1:1 Hex/EtOAc) to yield 62 mg (20%) of **50**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.66

(s, 1H), 7.68 (s, 1H), 7.44 (d, J = 8 Hz, 1H), 7.21 (t, J = 8 Hz, 1H), 6.94 (d, J = 8 Hz, 1H), 2.95 (s, 4H), 0.88 (s, 9H), 0.29 (s, 6H).

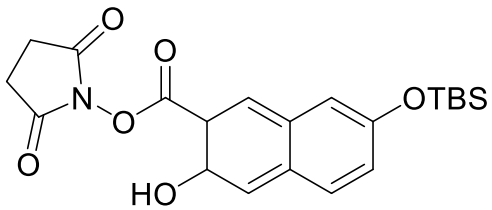
Preparation of 51[195]



51

3,7-Dihydroxy naphthoic-2-carboxylic acid (121 mg, 0.41 mmol) was azeotropically dried with pyridine (3 x 1 mL). TBDMSCl (155 mg, 1.02 mmol, 2.5 eq) was added to the flask. The contents were dissolved in DMF (4.1 mL, 100 mM). DBU (187 mg, 183 µL, 1.23 mmol, 3 eq) was added and the reaction was heated to 60 ºC. After 2 h, TLC (5% MeOH in DCM, R_f = 0.8) suggested complete conversion. The reaction was diluted with H_2O (5 mL) and the product was extracted with DCM (2 x 15 mL). The combined organic layers were washed with brine (1 x 30 mL), dried over Na_2SO_4, and concentrated. Column chromatography (DCM -> 1% MeOH in DCM) yielded 967 mg (81%) **51.** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.64 (d, J = 8 Hz, 1H), 7.24 (m, 4H), 1.11 (s, 18H), 0.47 (d, J = 8 Hz, 6H), 0.33 (s, 6H).

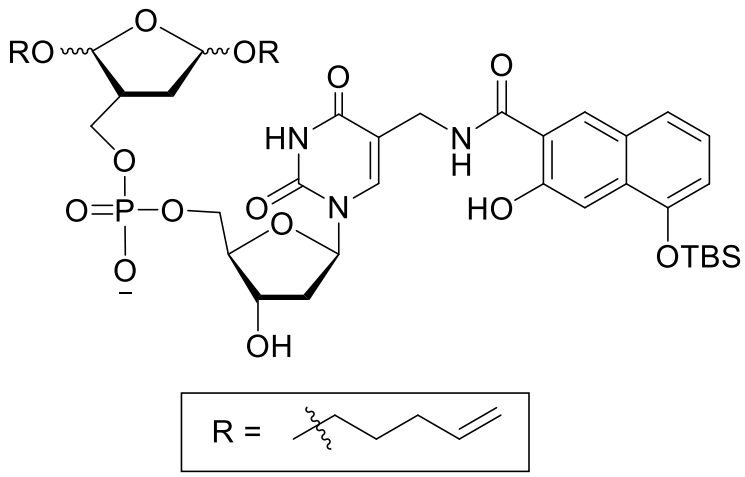
Preparation of 52[195]



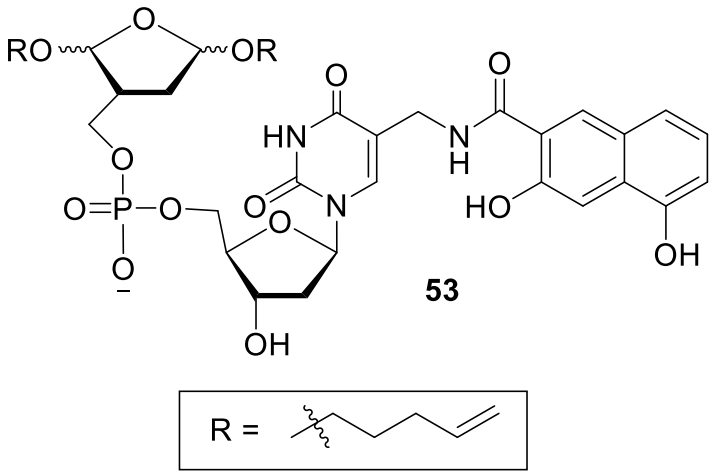
52

Compound **51** (86.5 mg, 0.2 mmol) was azeotropically dried with pyridine (1 x 0.8 mL). NHS (34.7 mg, 0.3 mmol, 1.5 eq) and EDC (57.8 mg, 0.3 mmol, 1.5 eq) were added to the flask and the reagents were dissolved in DMF and stirred overnight at 25 ºC. After 16 h, TLC (5% MeOH in DCM, R_f= 0.7) suggested the SM was completely converted so the reaction was concentrated. The product was purified by column chromatography (DCM -> 1% MeOH in DCM) to yield 62 mg (60%) of **52**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.93 (s, 1H), 8.55 (s, 1H), 7.59 (d, J = 8.8 Hz, 1H), 7.32 (s, 1H), 7.18 (d, J = 8.8 Hz, 1H), 2.94 (s, 4H), 1.01 (s, 9H), 0.24 (s, 6H).

Preparation of 53



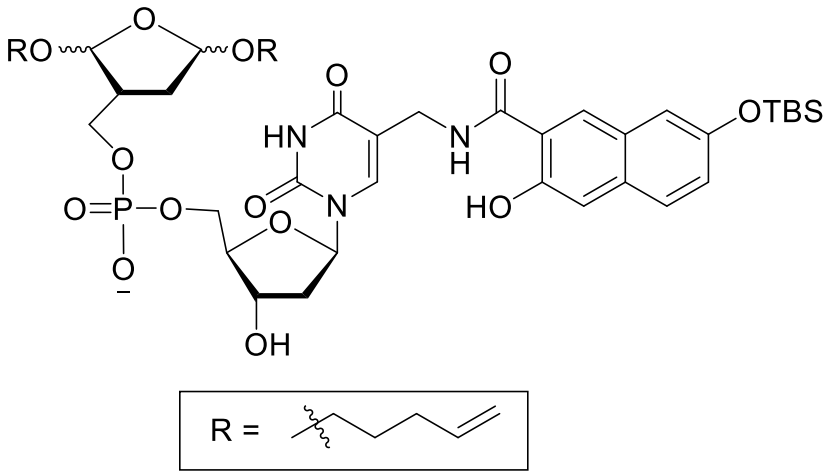
NHS ester **50** (10 mg, 16 μmol, 1.2 eq), scaffold **45** (8 mg, 14 μmol), and DIPEA (3 μL, 16 μmol, 1.2 eq) were dissolved in DMF (270 μL, 50 mM). The reaction was heated at 60 °C overnight. After 16 h, TLC (20% MeOH in DCM, R_f = 0.3) suggested the scaffold was still present but a significant number of new spots had formed so the reaction was concentrated and purified by column chromatography (15% MeOH in DCM -> 20% MeOH in DCM) to yield 11 mg of the amide intermediate (77%). The TBS group is labile and the NMR indicates these protons integrate to less than expected. Therefore, the product(s) isolate is likely a mixture of silylated and desilylated. This is not an issue because this intermediate is immediately desilylated (below). ^{1}H NMR (400 MHz, CD_3OD) δ 8.40 (m, 1H), 8.08 (s, 1H), 7.53 (d, *J* = 14 Hz, 1H), 7.13 (m, 1H), 6.84 (dd, *J* = 8, 14 Hz, 1H), 6.29 (m, 1H), 5.80 (m, 4H), 5.55 (m, 1H), 4.96 (m, 6H), 4.56 (m, 1H), 4.48 (m, 2H), 4.35 (m, 1H), 4.11 (m, 2H), 4.03 (m, 2H), 3.92 (m, 1H), 3.56 (m, 2H), 3.47 (m, 1H), 2.92 (m, 1H), 2.29 (m, 2H), 2.07 (m, 6H), 2.63 (m, 6H), 1.10 (s, 4H), 0.28 (s, 3H).



The amide intermediate (11 mg, 0.012 mmol) was dissolved in THF (120 µL, 100 mM). TEA•3HF (10 µL, 0.06 mmol, 5 eq) was added to the flask and the reaction was stirred at 25 °C overnight. After confirmed complete by TLC (15% MeOH in DCM, R_f = 0.4, UV active and stained with PAA), the reaction was concentrated and purified by column chromatography (10% MeOH in DCM -> 20% MeOH in DCM) provided 10 mg (99 %) of **53**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.44 (s, 1H), 8.02 (d, *J* = 8 Hz, 1H), 7.55 (s, 1H), 7.36 (d, *J* = 8 Hz, 1H), 7.10 (t, *J* = 8 Hz, 1H), 6.80 (d, *J* = 8 Hz, 1H), 6.29 (m, 1H), 5.79 (m, 4H), 4.94 (m, 6H), 4.49 (m, 5H), 4.30 (m, 1H), 4.08 (m, 4H), 3.57 (m, 2H), 2.30 (m, 2H), 2.11 (m, 3H), 2.08 (m, 4H), 1.60 (m, 4H).

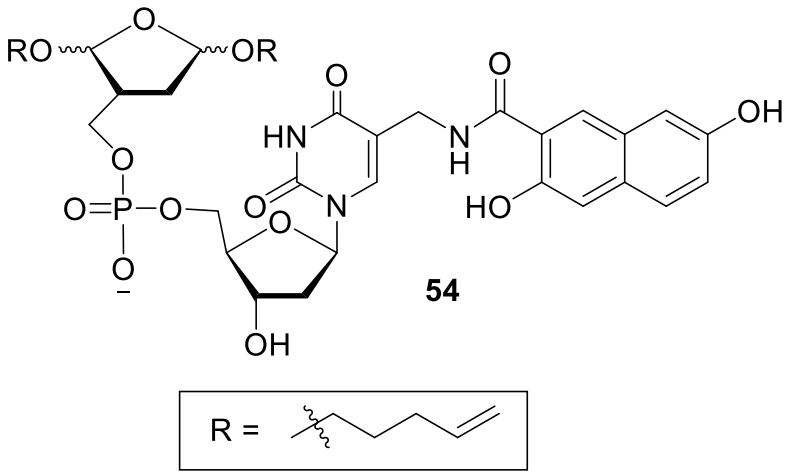
ESI-TOF m/z calculated for $C_{36}H_{45}N_3O_{14}P$ (M + H) – 777.2717, 777.2330 observed.

Preparation of 54



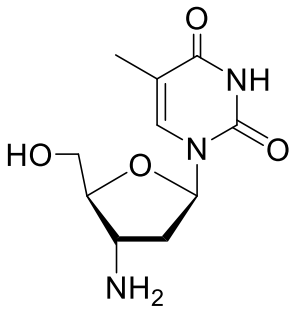
NHS ester **52** (10 mg, 16 µmol, 1.2 eq), scaffold **45** (8 mg, 14 µmol), and DIPEA (3 µL, 16 µmol, 1.2 eq) were dissolved in DMF (270 µL, 50 mM). The reaction was heated at 60 °C overnight. After 16 h, TLC (20% MeOH in DCM, R_f = 0.3) suggested the scaffold was still present but a significant number

of new spots had formed so the reaction was concentrated and purified by column chromatography (15% MeOH in DCM -> 20% MeOH in DCM) to yield 11 mg of the amide intermediate (77%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.67 (s, 2H), 7.96 (s, 1H), 7.86 (s, 1H), 7.55 (d, *J* = 9.6 Hz, 1H), 7.37 (d, *J* = 9.6 Hz, 1H), 7.25 (dd, *J* = 2.4, 8.8 Hz, 2H), 7.06 (m, 1H), 6.16 (t, *J* = 6.6 Hz, 1H), 5.80 (m, 2H), 4.94 (m, 4H), 4.45 (q, *J* = 2.4 Hz, 2H), 4.41 (s, 1H), 4.10 (d, *J* = 12 Hz, 2H), 4.05 (m, 1H), 3.89 (m, 1H), 3.85 (q, *J* = 3.3 Hz, 1H), 3.71 (dd, *J* = 3.3, 12 Hz, 1H), 3.65 (dd, *J* = 3.3 Hz, 4H), 2.23 (m, 4H), 2.18 (m, 3H), 1.56 (m, 4H), 0.99 (m, 5H), 0.88 (s, 7H), 0.08 (s, 4H).

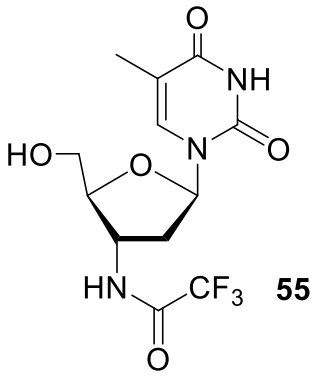


The amide intermediate (11 mg, 0.012 mmol) was dissolved in THF (120 µL, 100 mM). TEA•3HF (10 µL, 0.06 mmol, 5 eq) was added to the flask and the reaction was stirred at 25 °C overnight. After confirmed complete by TLC (15% MeOH in DCM, R_f = 0.4, UV active and stained with PAA), the reaction was concentrated and purified by column chromatography (10% MeOH in DCM -> 20% MeOH in DCM) provided 10 mg (99 %) of **54**. ^{1}H NMR (400 MHz, CD_3OD) δ 10.00 (s, 1H), 8.60 (s, 1H), 7.95 (s, 1H), 7.53 (d, *J* = 6 Hz, 1H), 7.34 (s, 1H), 7.16 (s, 1H), 7.09 (d, *J* = 6 Hz, 1H), 6.22 (t, *J* = 3.8 Hz, 1H), 5.84 (m, 4H), 4.99 (m, 5H), 4.46 (m, 4H), 4.07 (m, 4H), 3.92 (s, 1H), 3.77 (d, *J* = 11 Hz, 1H), 3.48 (m, 2H), 2.25 (m, 4H), 2.16 (m, 2H), 2.04 (m, 4H), 1.56 (m, 2H).

Preparation of 55[196,197]



AZT (2.19, 8.2 mmol) was dissolved in 50% MeOH, 30% tBuOH, 20% H_2O (82 mL, 100 mM). Activated palladium on carbon (1.3 g, 60 wt%) was added and flushed with a hydrogen balloon three times. The flask was sparged with H_2 for several hours. After 3 h, TLC (3% MeOH in DCM, UV active and stained with PAA) confirmed the starting material ($R_f = 0.8$) converted to the product ($R_f = 0.02$). The crude mixture was filtered through Celite and concentrated to yield 1.7 g of (86%) of the 3'-amine. ^{1}H NMR (400 MHz, CD_3OD) δ 7.87 (s, 1H), 6.18 (dd, *J* = 2, 6.8 Hz, 1H), 3.86 (dd, *J* = 2, 10 Hz, 1H), 3.77 (dd, *J* = 3.2, 10 Hz, 1H), 3.70 (quint, *J* = 3.2 Hz, 1H), 3.54 (q, *J* = 6.8 Hz, 1H), 2.26 (m, 1H), 2.20 (m, 1H), 1.88 (s, 3H).



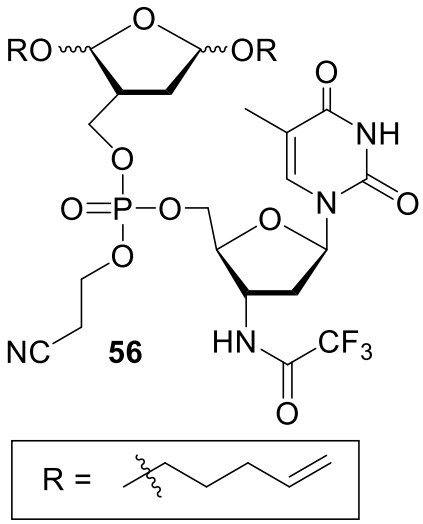
55

The crude mixture (1.72 g, 7.12 mmol) was dissolved in THF (25 mL, 280 mM). Triethylamine (14 mL, 106.8 mmol, 15 eq) and ethyl trifluoroacetate (8.4 mL, 71.3 mmol, 10 eq) were added to the flask to make a final concentration of 100 mM. The reaction was stirred for 4 h at 25 °C. TLC (3% MeOH in DCM, $R_f = 0.5$, UV active and stained by PAA) confirmed conversion to product. The reaction was concentrated and purified by column chromatography (1% MeOH in DCM to yield 2.29 g (93%) of **55**. ^{1}H NMR (400 MHz, CD_3OD) δ 7.85 (s, 1H), 6.27 (t, *J* = 6.4 Hz, 1H), 4.60 (q, *J* = 6.4 Hz, 1H), 3.97 (quint, *J* = 2.8 Hz, 1H), 3.85 (dd, *J* = 2.8, 12 Hz, 1H), 3.73 (dd, *J* = 2.8, 12 Hz, 1H), 2.40 (m, 2H), 1.89 (s, 3H).

^{13}C NMR (400 MHz, CD_3OD) δ 166.2, 159.4, 158.9, 158.6, 152.2, 138.0, 118.6, 115.8, 111.5, 85.9, 85.6, 62.3, 50.7, 37.9, 12.3, 9.09.

ESI-TOF m/z calculated for $C_{12}H_{14}F_3N_3O_5$ (M + H) − 338.0886, 338.3413 observed.

Preparation of 56

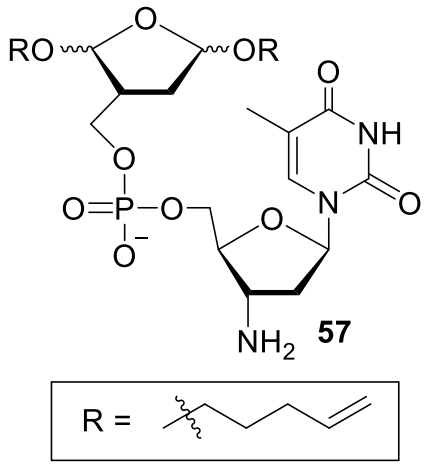


Compounds **10** (101 mg, 0.38 mmol, 1.2 eq) and **55** (172 mg, 0.32 mmol, 1 eq) were azeotropically dried together with toluene (2 x 1 mL). S-Ethyl tetrazole/MeCN (250 mM, 1.5 mL, 0.38 mmol, 1.2 eq) was added to the flask. After 4 h, TLC (1:1 EtOAc/Hex, R_f = 0.6, UV active and stained with PAA) indicated that the majority of **55** was consumed. tBuOOH (500 mM, 200 µL, 1 mmol, 3 eq) was added and the reaction was stirred for 15-20 min. The reaction was concentrated under vacuum and column chromatography (2:1 EtOAc/Hex) yielded 131 mg (56 %) of **56.** ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.36 (d, *J* = 10.1 Hz, 1H), 8.86 (d, *J* = 5 Hz, 1H), 7.46 (d, *J* = 2 Hz, 1H), 6.38 (q, *J* = 5 Hz, 1H), 5.74 (dsextet, *J* = 2, 4 Hz, 2H), 5.10-4.90 (m, 6H), 4.51 (s, 1H), 4.34 (s, 2H), 4.25 (q, *J* = 6 Hz, 2H), 4.16 (s, 2H), 4.06 (m, 1H), 3.63 (sextet, *J* = 2 Hz, 2H), 3.33 (m, 2H), 2.75 (q, *J* = 6 Hz, 3H), 2.44 (q, *J* = 8 Hz, 2H), 2.28 (m, 1H), 2.15 (m, 1H), 2.01 (d, *J* = 4 Hz, 4H), 1.96 (s, 3H), 1.58 (q, *J* = 4 Hz, 4H).

^{31}P NMR (400 MHz, $CDCl_3$) δ -1.97, -2.00. -2.04, -2.1, -2.4, -2.5, -2.6, -2.7, -2.9.

ESI-TOF m/z calculated for $C_{30}H_{42}F_3N_4O_{11}P$ (M + H) – 723.2540.2717, 723.2303 observed.

Preparation of 57



Compound **56** (115 mg, 0.15 mmol) was dissolved in concentrated aqueous ammonia (1.5 mL). The reaction was stirred at 0 °C for 4 h. After 4 h, TLC (20% MeOH in DCM, R_f = 0.2, UV active and stained with $KMnO_4$) confirmed the disappearance of starting material. The reaction was concentrated to yield 85 mg (100%) of **57**. ^{1}H NMR (400 MHz, CD_3OD) δ 6.34 (t, *J* = 5.8 Hz, 1H), 5.78 (m, 2H), 5.10-4.98 (m, 6H), 4.13 (s, 3H), 3.88 (m, 2H), 3.66 (m, 2H), 3.40 (q, *J* = 5.8 Hz, 2H), 2.81 (s, 2H), 2.68 (t, *J* = 7 Hz, 1H), 2.49 (m, 1H), 2.44 (m, 2H), 2.30 (m, 1H), 2.06 (t, *J* = 7 Hz, 4H), 1.93 (s, 3H), 1.60 (quint, *J* = 7 Hz, 4H).

^{31}P NMR (400 MHz, CD_3OD) δ 0.07.

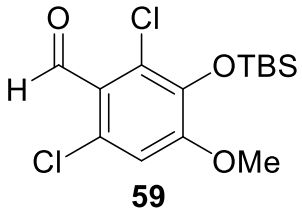
General Procedure for the preparation of Libraries 58 and 71

Amine scaffold **57** or **66** (100 nmol) was azeotropically dried with carboxylic acid (140 nmol, 1.4 eq) in pyridine (1 x 15 μL) using a Speed Vac concentrator in a 384-well microtiter plate (VWR). To each well, a 2X activating solution (5 μL; 28 mM HBTU and 28 mM HOBt in DMF), DIPEA (2 μL), and DMF (3 μL) were added. The final concentrations during reaction were: [**57**/**66**] = 10 mM, [acid] = 14 mM, [HBTU] = 14 mM, [HOBt] = 14 mM, 20% DIPEA in DMF. The well plate was shaken at 25 °C overnight. Some wells were analyzed by ESI-MS to confirm coupling efficiency. The solutions were concentrated using a Speed Vac concentrator and the well plate was covered and stored at -80 °C.

Immediately before an assay, the amide was thawed, dissolved in DMF (4 μL, 25 mM). An aliquot (2 μL, 50 nmol) was mixed with NBS (8 μL, 15 mM, 2.4 eq, 40% H_2O in MeCN) at 0

°C for 9 min. The concentrations during reaction were: [SM] = 5 mM, [NBS] = 12 mM, 20% DMF, 30% H_2O in MeCN. After 9 min, $Na_2S_2O_3$ (5 µL, 200 mM) was added and reaction quenched on ice for 10 min. Samples were concentrated with a Speed Vac concentrator. Some samples were analyzed ESI-MS to confirm product formation.

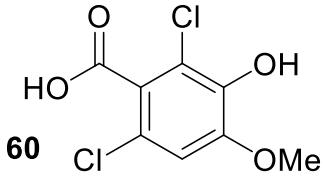
Preparation of 59[27]



59

Commercially available 2,6-dichloro-3-hydroxy-4-methoxy benzaldehyde (14 mg, 0.06 mmol) was azeotropically dried in pyridine (2 x 0.2 mL). TBSCl (12 mg, 0.08 mmol, 1.2 eq) and DMAP (5 mg, 0.02 mmol, 0.4 eq) were added. The reagents were dissolved in TEA (40 mg, 60 µL, 0.4 mmol, 7 eq) and DCM (450 µL, 150 mM). The reaction was stirred overnight at 25 °C. After 20 h, TLC (3% MeOH in DCM, R_f = 0.8, UV active) suggested reaction complete. The reaction was diluted with DCM (5 mL) and washed with bicarbonate solution (1 x 5 mL) and brine (1 x 5 mL). The combined aqueous layers were extracted with DCM (1 x 10 mL). The combined organic layers were dried over Na_2SO_4, concentrated, and purified by column chromatography (1:10 EtOAc/Hex -> 1:4 EtOAc/Hex) to yield 17 mg (85%) of **59**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.38 (s, 1H), 6.82 (s, 1H), 3.86 (s, 3H), 1.01 (s, 9 H), 0.18 (s, 6H).

Preparation of 60[164]

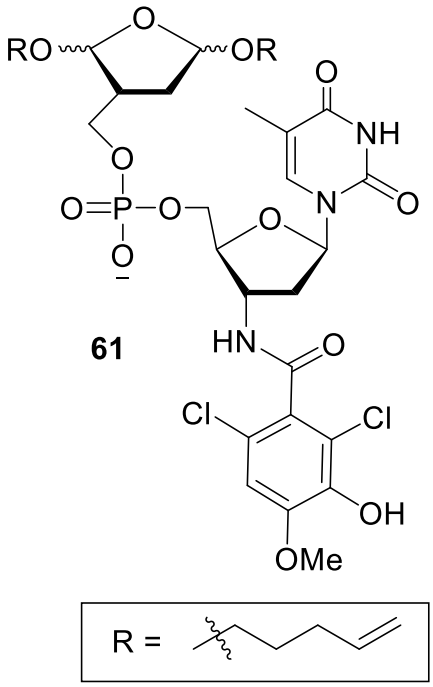


60

Aldehyde **59** (17.1 mg, 0.05 mmol) was dissolved in THF (0.18 mL). NaH_2PO_4 (6.8 mg, 0.03 mmol, 0.6 eq) and H_2O (71 µL) were added and the reaction stirred at 25 °C for 10 min. After 10 min, $NaClO_2$ (20 mg, 0.2 mmol, 4 eq) and 30% H_2O_2 (50 µL) were added and reaction was stirred at 25 °C. After 2.5 h, TLC (4:1 Hex/EtOAc, R_f = 0.1, UV active) suggested the starting material was consumed so the mixture was diluted with EtOAc and washed with H_2O (2 x 1 mL). The combined aqueous

layers were extracted with EtOAc (4 x 2 mL). The combined organic layers were dried over Na_2SO_4 and concentrated. The crude was purified by column chromatography (5% MeOH in DCM -> 10% MeOH in DCM) to yield 8.9 mg (50%) **60**. ^{1}H NMR (400 MHz, CD_3OD) δ 6.82 (s, 1H), 3.92 (s, 3H).

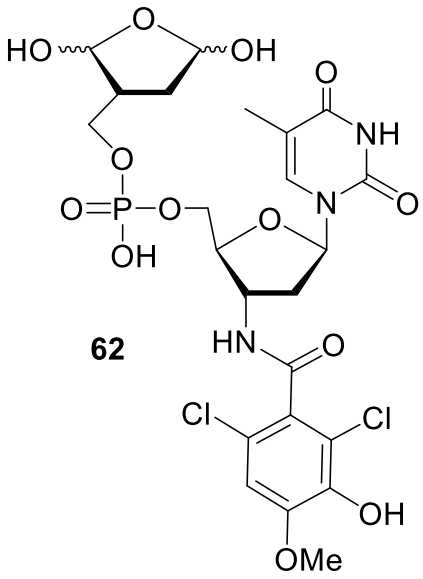
Preparation of 61



Scaffold **57** (10 mg, 0.02 mmol) and **60** (5.7 mg, 0.025 mmol, 1.4 eq) were azeotropically dried together with pyridine (2 x 0.5 mL). HBTU (9.3.0 mg, 0.025 mmol, 1.4 eq) and HOBt (9.3 mg, 0.025 mmol, 1.4 eq) were added to the starting materials. The contents were dissolved in DMF (350 mL) and DIPEA (70 µL). The reaction was stirred at 25 ^oC overnight. TLC (20% MeOH in DCM, R_f = 0.3, UV active and stained with PAA) suggested **60** was consumed. The reaction was concentrated and purified by column chromatography (5% MeOH in DCM) to yield 10.9 mg (79%) of **61**. ^{1}H NMR (400 MHz, CD_3OD) δ 8.05 (dd, *J* = 1.4, 8.5 Hz, 1H), 7.91 (t, *J* = 1.4 Hz, 1H), 7.51 (t, *J* = 8.5 Hz, 1H), 6.98 (s, 1H), 6.37 (m, 1H), 6.27 (m, 1H), 5.80 (m, 2H), 5.16-4.95 (m, 6H), 4.67 (m, 1H), 4.62 (m, 1H), 4.25 (m, 1H), 4.20 (m, 2H), 4.10 (m, 2H), 3.65 (m, 3H), 2.72 (m, 1H), 2.42 (m, 2H), 2.29 (m, 2H), 2.08 (m, 4H), 1.97 (s, 3H), 1.62 (t, *J* = 3.4 Hz, 4H).

^{31}P NMR (400 MHz, CD_3OD) δ -2.3.

Preparation of 62

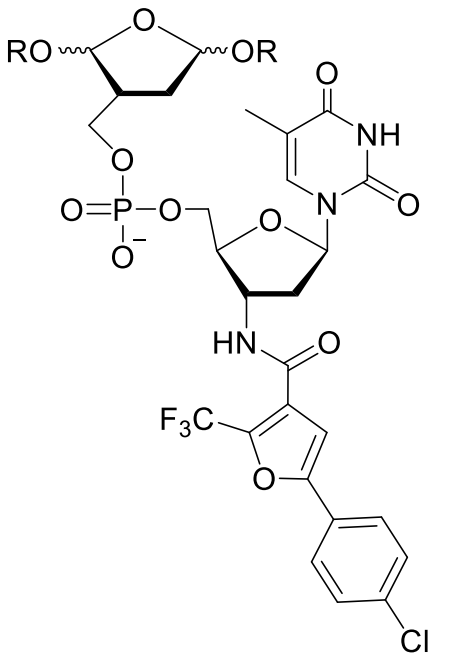


62

Compound **61** (2 μL, 20 mM, 40 nmol) was mixed with H_2O (3 μL) and NBS (5 μL, 25 mM, 2.5 eq, MeCN) at 0 °C for 9 min. The concentrations during reaction were: [SM] = 5 mM, [NBS] = 12.5 mM, 20% DMF, 30% H_2O in MeCN. After 9 min, $Na_2S_2O_3$ (10 μL, 200 mM) was added and reaction quenched on ice for 10 min. The crude material was concentrated by Speed Vacuum concentrator and analyzed by ESI-MS.

ESI-MS m/z calculated for $C_{23}H_{26}Cl_2N_3O_{12}P$ (M – H_2O+ H) – 636.06, 636.21 observed.

Preparation of 63



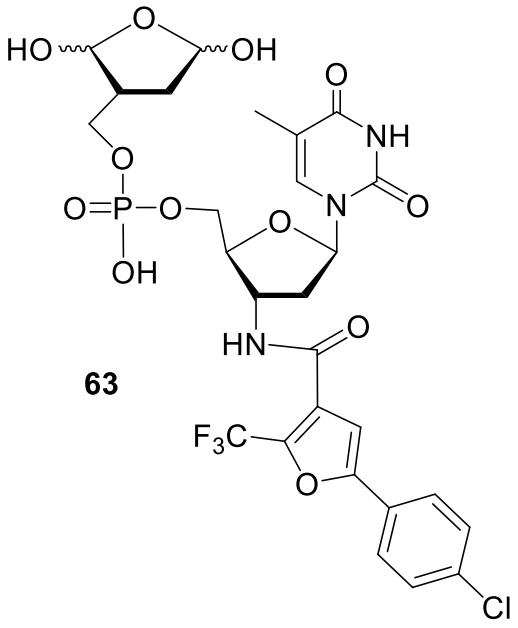
Scaffold **57** (17 mg, 0.03 mmol) and 5-(4-chlorophenyl)-3-(trifluoromethyl) furan-2-carboxylic acid (11 mg, 0.04 mmol, 1.4 eq) were azeotropically dried together with pyridine (2 x 1 mL). HBTU (14.0 mg, 0.037 mmol, 1.4 eq) and HOBt (5 mg, 0.037 mmol, 1.4 eq) were added to the starting materials. The contents were dissolved in DMF (1.04 mL) and DIPEA (260 μL). The reaction was stirred at 25 °C overnight. TLC (20% MeOH in DCM, R_f = 0.3, UV active and stained with PAA) suggested **57** was consumed. The reaction was concentrated and purified by column chromatography (5% MeOH in DCM) to yield 21.2 mg (96%) of the precursor to **63**. ^{1}H NMR (400 MHz, CD_3OD) δ 7.76 (d, *J* = 6.8 Hz, 1H), 7.70 (d, *J*

= 7.6 Hz, 2H), 7.59 (d, *J* = 6.8 Hz, 2H), 7. 50 (d, *J* = 7.6 Hz, 1H), 7.35 (quint, *J* = 6.8 Hz, 4H), 6.37 (t, *J* = 3.6 Hz, 1H), 5.78 (m, 2H), 5.00 (m, 6H), 4.68 (m, 1H), 4.15 (d, *J* = 1.6 Hz, 1H), 4.10 (m, 2H), 3.90 (m, 2H), 3.66 (dd, *J* = 1.6, 5.2 Hz, 2H), 2.49 (m, 1H), 2.40 (d, *J* = 1.6 Hz, 2H), 2.30 (m, 1H), 2.07 (d, *J* = 6.4 Hz, 4H), 1.97 (s, 3H), 1.90 (m, 1H), 1.60 (quint, *J* = 6.4 Hz, 4H).

^{31}P NMR (400 MHz, CD_3OD) δ -2.4, -3.0.

ESI-TOF m/z calculated for $C_{37}H_{44}ClF_3N_3O_{12}P$ (M - H) − 844.2303, 844.2221 observed.

$\varepsilon_{273} = 2.0 \times 10^4\ M^{-1}cm^{-1}$



63

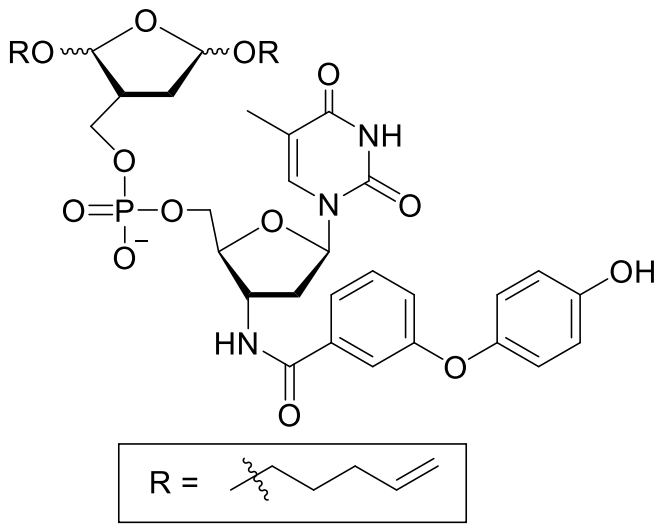
The pentenoyl protected inhibitor was stored as a 25 mM stock solution in DMF at -20 ºC. An aliquot of the starting material (2 µL, 50 nmol) was placed in a microtiter well plate and cooled to 4 ºC. Cold H_2O (3 µL) was added to the well. NBS in MeCN (25 mM, 5 µL) was added and the well plate was shaken at 4 ºC for 9 min. After 9 min, the reaction was quenched by an equal volume of $Na_2S_2O_3$ (200 mM in H_2O, 10 µL) and the plate was shaken for additional 5-10 min. The reaction was concentrated by speed vacuum, redissolved in 1:1 MeCN/H_2O and used directly.

Reaction Conditions: [SM] = 5 mM, [NBS] = 12.5 mM, 30% H_2O, 20% DMF, 50% MeCN

Quenching Conditions: [SM] = 2.5 mM, [NBS] = 6.25 mM, [$Na_2S_2O_3$] = 100 mM, 65% H_2O, 10% DMF, 25% MeCN.

ESI-MS m/z calculated for $C_{27}H_{27}ClF_3N_3O_{12}P$ (M - H) – 708.09, 708.20 observed.

Preparation of 64

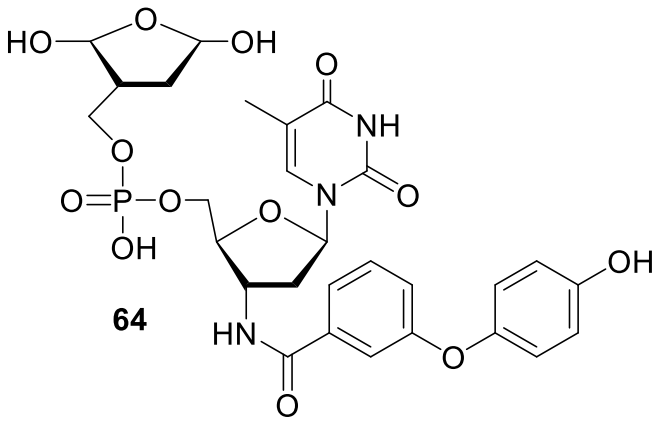


Scaffold **57** (15 mg, 0.03 mmol) and 3-(4-hydroxyphenoxy) benzoic acid (8.5 mg, 0.04 mmol, 1.4 eq) were azeotropically dried together with pyridine (2 x 1 mL). HBTU (14.0 mg, 0.04 mmol, 1.4 eq) and HOBt (5 mg, 0.04 mmol, 1.4 eq) were added to the starting materials. The contents were dissolved in DMF (1.04 mL) and DIPEA (260 µL). The reaction was stirred at 25 ºC overnight. TLC (20% MeOH in DCM, R_f = 0.3, UV active and stained with PAA) suggested **57** was consumed. The reaction was concentrated and purified by column chromatography (5% MeOH in DCM) to yield 9.4 mg (46%) of the precursor to **64**. ^{1}H NMR (400 MHz, CD_3OD) δ 7.89 (s, 2H), 7.81 (d, *J* = 8.8 Hz, 2H), 7.71 (s, 1H), 6.91 (dd, *J* = 3.2, 8.8 Hz, 2H), 6.90 (s, 1H), 6.83 (s, 2H), 6.38 (t, *J* = 3.2, 1H), 5.80 (m, 2H), 5.08-4.95 (m, 6H), 4.69 (m, 1H), 4.16 (s, 1H), 4. 23 (m, 2H), 4.10 (m, 2H), 3.97 (m, 1H), 3.88 (m, 2H), 3.63 (m, 2H), 2.42 (m, 3H), 2.07 (q, *J* = 1.6 Hz, 4H), 1.97 (s, 3H), 1.89 (m, 1H), 1.60 (q, *J* = 1.6 Hz, 4H).

^{31}P NMR (400 MHz, CD_3OD) δ -2.01.

ESI-TOF m/z calculated for $C_{38}H_{48}N_3O_{13}P$ (M - H) − 784.2925 calculated, 784.4690 and 767.2130 (-OH) observed.

$\varepsilon_{263} = 2.22 \times 10^4\ M^{-1}cm^{-1}$



64

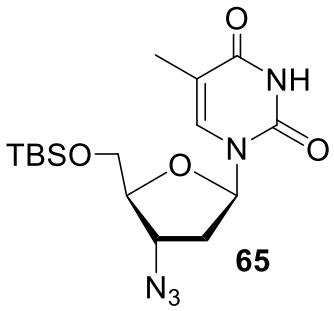
The pentenoyl protected inhibitor was stored as a 25 mM stock solution in DMF at -20 ºC. An aliquot of the starting material (2 µL, 50 nmol) was placed in a microtiter well plate and cooled to 4 ºC. Cold H_2O (3 µL) was added to the well. NBS in MeCN (25 mM, 5 µL) was added and the well plate was shaken at 4 ºC for 9 min. After 9 min, the reaction was quenched by an equal volume of $Na_2S_2O_3$ (200 mM in H_2O) and the plate was shaken for an additional 5-10 min. The reaction was concentrated by speed vacuum, redissolved in 1:1 MeCN/H_2O and used directly.

Reaction Conditions: [SM] = 5 mM, [NBS] = 12.5 mM, 30% H_2O, 20% DMF, 50% MeCN

Quenching Conditions: [SM] = 2.5 mM, [NBS] = 6.25 mM, [$Na_2S_2O_3$] = 100 mM, 65% H_2O, 10% DMF, 25% MeCN.

ESI-MS m/z calculated for $C_{28}H_{31}N_3O_{13}P$ (M - H) – 648.16, 648.65 observed.

Preparation of 65[198]



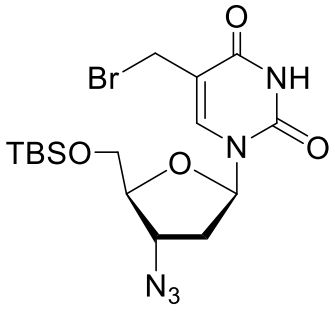
65

AZT (587 mg, 2.19 mmol, 1 eq) was azeotropically dried twice with pyridine (2 x 1 mL). Imidazole (604 mg, 8.76 mmol, 4 eq) and TBDMSCl (672 mg, 4.38 mmol, 2 eq) were added and dissolved in DMF (4.5 mL, 500 mM). The reaction was heated to 50° C overnight. When confirmed complete by TLC (5% MeOH in DCM, R_f = 0.5, UV active and stained with PAA), the reaction was cooled and diluted with EtOAc (10 mL). The organic layer was washed with sat. NH_4Cl (3 x 15 mL) and brine (2 x 15 mL). The combined aqueous layers were extracted using EtOAc (2 x 50 mL). The combined organic layers were dried

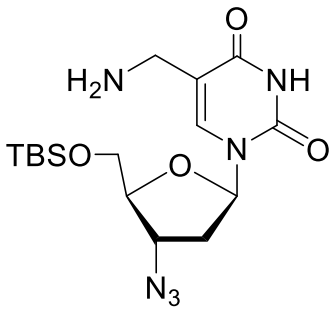
over Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (2:1 Hex/EtOAc) yielding 825 mg (95%) of **65**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.72 (s, 1H), 7.43 (s, 1H), 6.22 (t, *J* = 6.8 Hz, 1H), 4.24 (q, *J* = 3.2 Hz, 1H), 3.96 (d, *J* = 3.2 Hz, 1H), 3.94 (dd, *J* = 2.8, 14 Hz, 1H), 3.79 (dd, *J* = 2.8, 14 Hz, 1H), 2.44 (m, 1H), 2.22 (m, 1H), 1.91 (s, 3H), 0.93 (s, 9H), 0.13 (s, 6H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 134.9, 111.0, 84.5, 84.4, 62.9, 60.5, 37.9, 25.9, 25.7, 25.6, 18.3, 12.5, -5.4, -5.5.

Preparation of 66



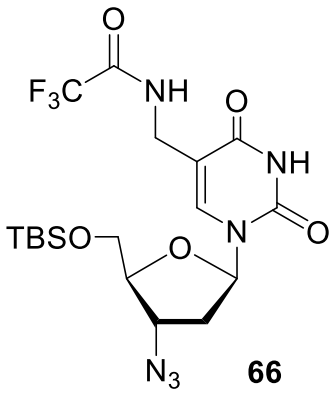
Compound **65** (982.8 mg, 2.58 mmol, 1 eq) was mixed with NBS (366 mg, 2.06 mmol, 0.8 eq) and dissolved in distilled benzene (13 mL, 200 mM). The mixture was sparged with Ar for 20 minutes. The reaction was activated by sun lamp. The reaction was stirred for 1 h while the mixture turned red. When TLC (1:1 Hex/EtOAc, UV active and stained with iodine) showed a mixture of starting material and one new spot, a second aliquot of 0.8 eq NBS was added to the reaction. After a total of 2 h and 2 x 0.8 eq NBS added, the SM was consumed to yield one major spot. The mixture was immediately filtered through a glass frit and diluted with DCM (10 mL) and H_2O (20 mL). The aqueous layer was extracted with DCM (2 x 20 mL) and the combined organic layers were washed with sat. bicarbonate solution (1 x 40 mL), brine (1 x 40 mL), dried over Na_2SO_4, and concentrated under vacuum to give 853.6 mg (72%) a yellow crude residue. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.67 (s, 1H), 7.84 (s, 1H), 6.17 (t, *J* = 6.4 Hz, 1H), 4.24 (m, 2H), 4.20 (m, 1H), 3.98 (m, 1H), 3.96 (dd, *J* = 2.8, 11.6 Hz, 1H), 3.80 (dd, *J* = 2.8, 11.6 Hz, 1H), 2.50 (m, 1H), 2.25 (m, 1H), 0.94 (s, 9H), 0.14 (s, 6H).



The crude material (853.2 mg, 1.85 mmol) was dissolved in EtOH (13 mL). Concentrated aqueous ammonia (23 mL, 15.7 M, 47.1 mmol, 25 eq) was added to the flask. The reaction was stirred at 25 °C for 1.5 h. After 1.5 h, TLC (10% MeOH in DCM, R_f = 0.3) suggested the bromide intermediate was consumed. The reaction was concentrated and purified by column chromatography (DCM -> 3% MeOH in DCM) to yield 339 mg (43%) of the second intermediate. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.78 (s, 1H), 6.06 (t, *J* = 6.4 Hz, 1H), 4.22 (m, 1H), 3.88 (q, *J* = 4.4 Hz, 1H), 3.82 (d, *J* = 4.4 Hz, 1H), 3.78 (d, *J* = 4.4 Hz, 1H), 3.73 (s, 2H), 2.35 (t, *J* = 6.4 Hz, 2H), 0.85 (s, 9H), 0.05 (s, 6H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 164.3, 150.3, 140.2, 109.2, 85.1, 84.6, 62.9, 60.6, 50.1, 37.3, 25.8, 22.4, 18.3, -5.4, -5.5.

ESI-TOF m/z calculated for $C_{16}H_{28}N_6O_4Si$ (M + H) – 397.1941 calculated, 397.2012 observed.



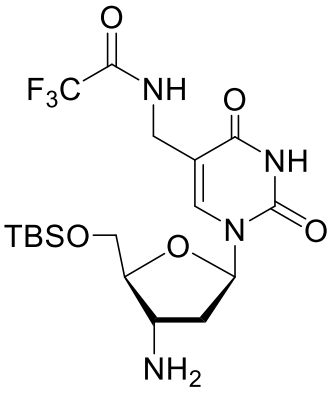
66

The second intermediate (313.5 mg, 0.791 mmol) was dissolved in THF (5.2 mL, 280 mM). Triethylamine (1.6 mL, 11.8 mmol, 15 eq) and ethyl trifluoroacetate (0.95 mL, 7.91 mmol, 10 eq) were added to the flask to make the final concentration 100 mM. The reaction was stirred for 4 hours at 25 °C. TLC (3% MeOH in DCM, R_f = 0.5, UV active and stained by PAA) confirmed conversion to product. The reaction was concentrated and carried further to the next step without further purification, yielding 500 mg of **66** (91%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.17 (s, 1H), 7.91 (s, 1H), 7.80 (s, 1H), 6.13 (t, *J* = 6.4 Hz, 1H), 4.23 (q, *J* = 4 Hz, 1H), 4.20 (s, 2H), 4.00 (q, *J* = 4 Hz, 1H), 3.91 (dd, *J* = 3.6, 11.6 Hz, 1H), 3.84 (dd, *J* = 3.6, 11.6 Hz, 1H), 2.49 (m, 1H), 2.24 (m, 1H), 0.91 (s, 9H), 0.13 (s, 6H).

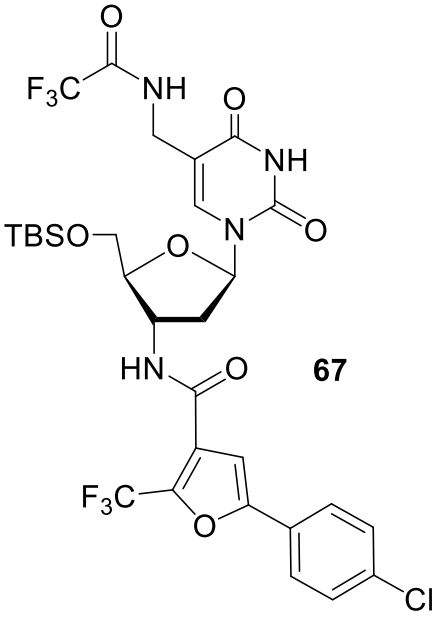
^{13}C NMR (400 MHz, $CDCl_3$) δ 162.9, 161.3, 160.9, 149.9, 138.7, 117.8, 114.9, 109.2, 84.8, 84.6, 60.8, 59.9, 53.2, 49.2, 45.5, 37.4, 35.7, 7.9.

ESI-TOF m/z calculated for $C_{18}H_{27}F_3N_6O_5Pi$ (M + H) – 493.1764, 493.1716 observed.

Preparation of 67



Compound **66** (22.5 mg, 0.05 mmol) was dissolved in 50% MeOH, 20% H_2O, 30% tBuOH (1 mL). Activated Pd/C (14.1 mg, 60 wt%) was added to the solution. The reaction was continuously sparged with H_2 (1 atm). After 1 h, TLC (5% MeOH in DCM, UV active and stained with PAA, R_f = 0.35) suggested the reaction was complete. The mixture was passed through a Celite column and concentrated to 17 mg (82%) of crude amine, which was carried forward without further purification. ^{1}H NMR (400 MHz, CD_3OD) δ 7.83 (s, 1H), 6.18 (t, *J* = 6.4 Hz, 1H), 4.12 (s, 2H), 3.93 (dd, *J* = 4, 11.2 Hz, 1H), 3.88 (dd, *J* = 4, 11.2 Hz, 1H), 3.77 (dt, *J* = 1.2, 4 Hz, 1H), 3.50 (q, *J* = 6.4 Hz, 1H), 2.25 (m, 2H), 0.93 (s, 9H), 0.12 (d, *J* = 1.2 Hz, 6H).



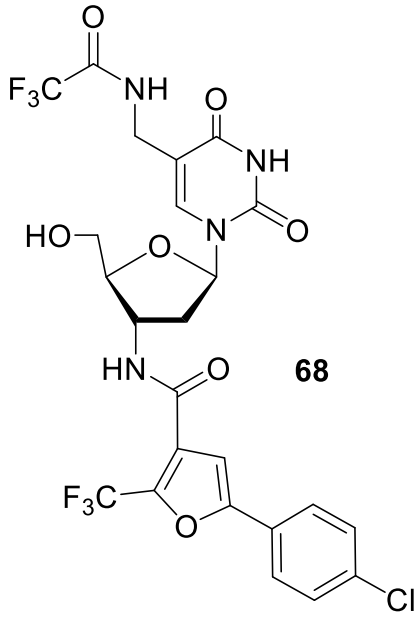
67

Crude amine (17 mg, 0.033 mmol) and 5-(4-chlorophenyl)-3-(trifluoromethyl) furan-2-carboxylic acid (14.2 mg, 0.05 mmol, 1.4 eq) were azeotropically dried together with pyridine (2 x 1 mL). HBTU (17.8 mg, 0.05 mmol, 1.4 eq) and HOBt (6.3 mg, 0.05 mmol, 1.4 eq) were added to the starting materials. The contents were dissolved in DMF (528 µL) and DIPEA (132 µL). The reaction was stirred at 25 °C overnight. TLC (10% MeOH in DCM, R_f = 0.4, UV active and stained with PAA) suggested the starting material was gone. The

reaction was concentrated and purified by column chromatography (1:1 Hex/EtOAc -> 5% MeOH, 1:1 Hex/EtOAc) to yield 18.1 mg (75%) of **67**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.23 (s, 1H), 8.10 (d, *J* = 6.4 Hz, 1H), 7.86 (s, 1H), 7.75 (d, *J* = 8.2 Hz, 2H), 7.63 (d, *J* = 8.2 Hz, 2H), 6.09 (t, *J* = 6.4 Hz, 1H), 4.59 (t, *J* = 4 Hz, 1H), 4.12 (t, *J* = 4 Hz, 2H), 3.88 (m, 2H), 2.43 (m, 1H), 2.25 (m, 1H), 0.82 (s, 9H), 0.04 (s, 6H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 162.7, 160.8, 150.1, 134.9, 129.0, 126.5, 125.8, 125.6, 120.1, 117.3, 110.7, 108.9, 105.9, 85.6, 54.7, 38.4, 36.5, 31.4, 25.7, 18.2, 16.9, -5.6. -5.7.

Preparation of 68



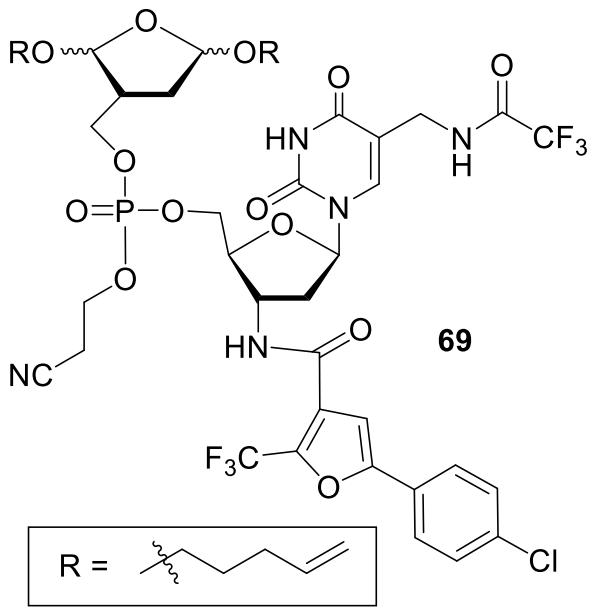
68

Compound **67** (135 mg, 0.19 mmol) was dissolved in THF (3.6 mL). TEA•3HF (74 mg, 0.46 mmol, 0.08 mL, 2.5 eq) was added to the flask. The reaction was stirred at 25 ºC overnight. TLC (5% MeOH in DCM, R_f = 0.5, UV active and stained with PAA) confirmed conversion to product. The reaction was concentrated and purified by column chromatography (DCM -> 2% MeOH) to yield 100 mg (86%) of **68**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.28 (t, *J* = 6.9, 1H), 7.79 (m, 3H), 7.74 (dd, *J* = 1.6, 6.9 Hz, 1H), 7.57 (dd, *J* = 1.6, 8 Hz, 1H), 7.23 (d, *J* = 8 Hz, 1H), 7.19 (m, 2H), 6.06 (q, *J* = 4.4 Hz, 1H), 4.54 (t, *J* = 7.6 Hz, 1H), 4.32 (d, *J* = 7.6 Hz, 1H), 4.21 (dd, *J* = 5.6, 14.8 Hz, 1H), 4.15 (dd, *J* = 5.6, 14.8 Hz, 1H), 3.84 (d, *J* = 10.8 Hz, 1H), 3.74 (d, *J* = 10.8 Hz, 1H), 2.30 (m, 2H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 163.2, 160.8, 157.2, 153.0, 149.9, 134.1, 128.6, 127.4, 126.1, 124.4, 123.9, 120.2, 117.9, 110.9, 108.6, 107.4, 83.9, 50.1, 46.1, 8.31, 8.29, 8.26, 8.24, 8.23.

ESI-TOF m/z calculated for $C_{24}H_{19}ClF_3N_4O_7$ (M + H) − 625.0846, 625.0844 observed.

Preparation of 69



Compounds **68** (61 mg, 0.08 mmol, 1.2 eq) and **10** (32 mg, 0.07 mmol, 1 eq) were azeotropically dried with toluene (2 x 0.5 mL). S-Ethyl tetrazole/MeCN (250 mM, 1 mL, 0.24 mmol, 1.2 eq) was added to the reaction flask. After 3 h, TLC (2:1 EtOAc/Hex, R_f = 0.5, UV active and stained with PAA) indicated that the majority of **68** was consumed. tBuOOH (500 mM, 18 mg, 40 µL, 0.20 mmol, 3 eq) was added for 15-20 min. The reaction was concentrated and column chromatography (2:1 EtOAc/Hex) yielded 61 mg (90 %) of **69**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.87 (d, *J* = 5.9, 1H), 7.56 (d, *J* = 8.4 Hz, 2H), 7.35 (d, *J* = 8.4 Hz, 2H), 7.07 (d, *J* = 3.6 Hz, 1H), 6.22 (t, *J* = 5.9, 1H), 5.78 (m, 2H), 4.99 (m, 6H), 4.75 (m, 1H), 4.40 (m, 3H), 4.30 (m, 4H), 4.28 (m, 2H), 3.88 (m, 2H), 3.61 (m, 2H), 3.38 (m, 2H), 2.78 (quint, *J* = 3.6 Hz, 2H), 2.63 (m, 1H), 2.52 (m, 2H), 2.45 (m, 1H), 2.08 (m, 4H), 1.88 (m, 1H), 1.62 (m, 4H).

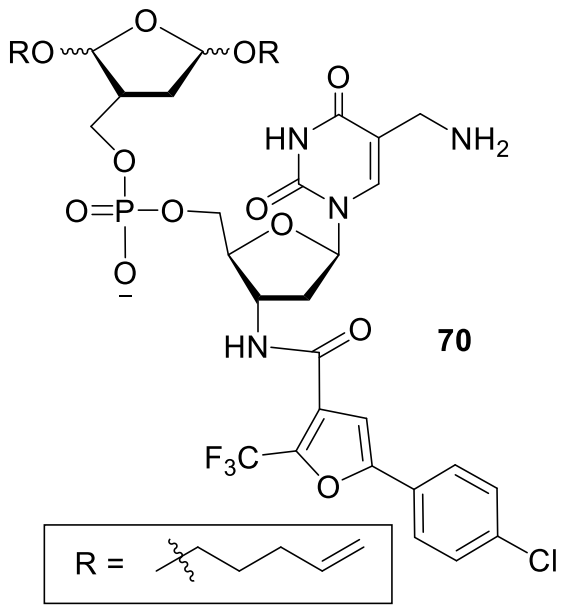
^{13}C NMR (400 MHz, $CDCl_3$) δ 137.9, 135.1, 129.0, 126.5, 125.6, 114.6, 114.5, 66.6, 65.8, 30.1, 30.0, 28.6, 269, 26.1, 14.6.

^{31}P NMR (400 MHz, $CDCl_3$) δ 7.8, -0.9, -1.0, -2.57, -2.62, -2.63, -2.65, -2.67, -2.69, -2.71, -3.06, -3.1.

^{19}F NMR (300 MHz, $CDCl_3$) δ -64.02, -64.03, -64.1, -78.12, -78.13 -78.2.

ESI-TOF m/z calculated for $C_{42}H_{47}ClF_6N_5O_{13}P$ (M + H) − 1010.2501, 1010.2183 observed.

Preparation of 70



70

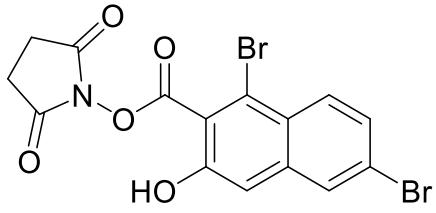
Compound **67** (8 mg, 0.008 mmol) was dissolved in concentrated aqueous ammonia (160 µL). The reaction was capped and stirred at 25 °C. After 5 h, TLC (15% MeOH in DCM, R_f = 0.1, UV active and stained with PAA) confirmed the disappearance of starting material and formation of a new spot. The reaction was concentrated and purified by column chromatography (5% MeOH in DCM -> 20% MeOH in DCM) to yield 9.3 mg (73%) of **70**. ^{1}H NMR (400 MHz, CD_3OD) δ 8.37 (s, 1H), 7.76 (d, *J* = 11.6 Hz, 2H), 7.68 (q, *J* = 5.2 Hz, 1H), 7.50 (d, *J* = 11.6 Hz, 2H), 7.24 (m, 3H), 6.28 (t, *J* = 5.2 Hz, 1H), 5.80 (m, 2H), 4.95 (m, 6H), 4.78 (q, *J* = 8.6 Hz, 1H), 4.26 (m, 2H), 4.08 (m, 2H), 3.83 (m, 4H), 3.63 (m, 2H), 2.53 (q, *J* = 8.6 Hz, 2H), 2.47 (m, 2H), 2.09 (m, 4H), 1.60 (m, 4H).

^{13}C NMR (400 MHz, CD_3OD) δ 166.3, 164.8, 139.4, 137.9, 136.3, 130.3, 128.4, 128.3, 127.0, 115.0, 114.98, 111.3, 108.1, 106.0, 68.5, 68.2, 62.2, 55.7, 43.7, 36.8, 35.1, 31.5, 31.4, 31.39, 31.35, 30.03, 30.00, 19.2, 18.6, 17.2, 13.0, 12.5.

^{31}P NMR (400 MHz, CD_3CN) δ -2.3.

ESI-TOF m/z calculated for $C_{37}H_{45}ClF_3N_4O_{12}P$ (M - H) − 859.2412, 859.2277 observed.

Preparation of NHS-Ester for 72



X

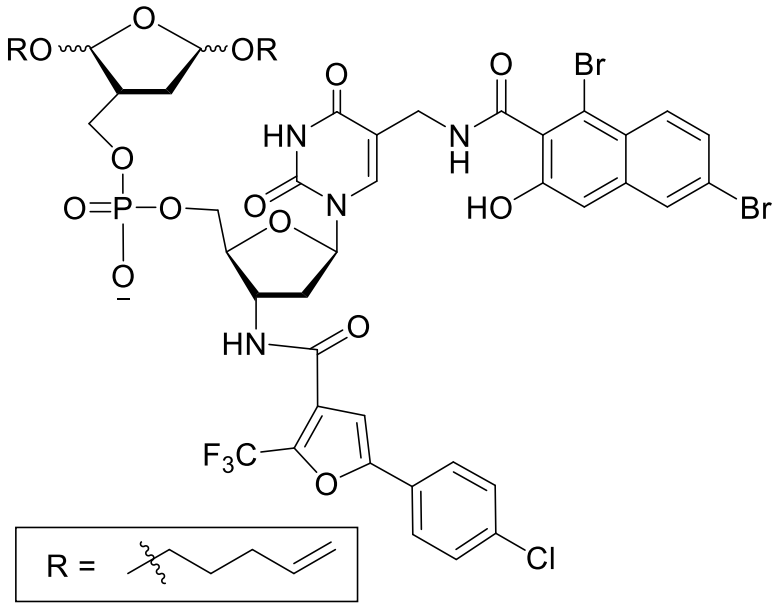
Commercially available 1,6-dibromo-3-hydroxy-2-naphthoic acid (102 mg, 0.3 mmol) was azeotropically dried in pyridine (2 x 1 mL). NHS (54 mg, 0.45 mmol, 1.5 eq),

and EDC (86 mg, 0.45 mmol, 1.5 eq) were added to the flask and the reagents were dissolved in DMF and stirred overnight at 25 ºC. After 16 h, TLC (2% MeOH in DCM, R_f = 0.8, UV active) suggested the SM was completely converted so the reaction was concentrated. The product was purified by column chromatography (DCM -> 1% MeOH in DCM) to yield 119 mg (90%) of NHS ester (**X**). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.79 (s, 1H), 8.60 (s, 1H), 8.09 (d, *J* = 8.8 Hz, 1H), 8.01 (s, 1H), 7.75 (d, *J* = 8.8 Hz, 1H), 2.97 (s, 4H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 169.6, 162.5, 153.4, 134.6, 133.6, 131.2, 130.6, 128.2, 127.7, 118.3, 115.0, 107.2, 53.4, 53.1, 36.4, 31.4, 29.6.

ESI-TOF m/z calculated for $C_{15}H_9NO_5$ (M + H) − 441.8847, 441.3207 observed.

Preparation of 72



Scaffold **70** (14 mg, 0.016 mmol) and **X** (11 mg, 0.024 mmol, 1.5 eq) were azeotropically dried together with pyridine (2 x 1 mL). The contents were dissolved in DMF (350 µL) and DIPEA (5 µL). The reaction was stirred at 25 ºC overnight. TLC (20% MeOH in DCM, R_f = 0.3, UV active and stained with PAA) suggested the starting materials were converted to a new spot. The reaction was concentrated and purified by column chromatography (2:1 Hex/EtOAc -> 1:1 Hex/EtOAc -> 1:1 Hex/EtOAc, 5% MeOH) to yield 12.2 mg (45%) of the precursor to **72**. The product was passed through a DOWEX (Na^+) column for clearer NMR analysis. ^{1}H NMR (400 MHz, CD_3OD) δ 8.38

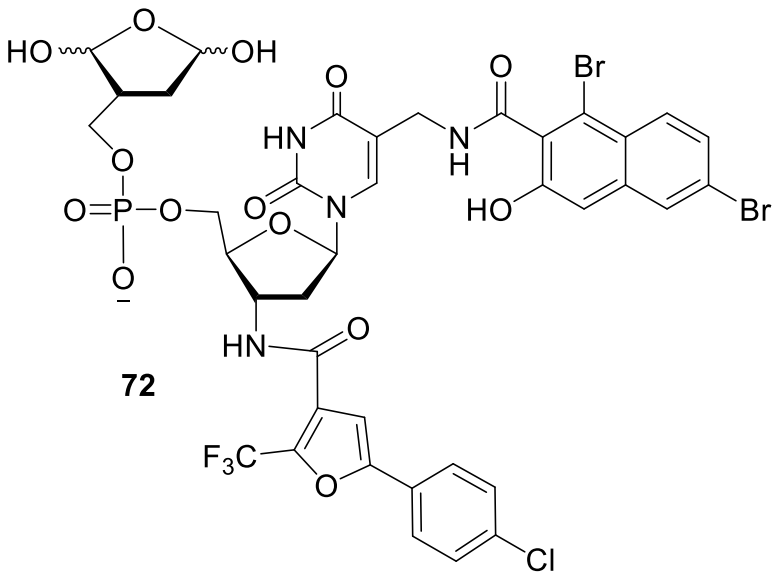
(s, 2H), 8.02 (s, 3H), 7.98 (d, *J* = 7.2 Hz, 2H), 7.74 (d, *J* = 8.8 Hz, 1H), 7.61 (d, *J* = 7.2, 2H), 7.48 (d, *J* = 8.8 Hz, 1H), 7.24 (m, 1H), 6.30 (s, 1H), 5.82 (m, 2H), 4.99 (m, 6H), 4.73 (s, 1H), 4.44 (s, 1H), 4.22 (s, 1H), 4.1 (m, 1H), 3.73 (m, 3H), 3.66 (m, 2H), 3.41 (m, 2H), 2.53 (m, 1H), 2.48 (m, 2H), 2.12 (q, *J* = 7.2 Hz, 4H), 2.05 (m, 2H), 1.63 (q, *J* = 7.2 Hz, 4H).

^{13}C NMR (400 MHz, CD_3OD) δ 164.8, 132.6, 132.2, 131.1, 129.9, 128.1, 117.6, 36.9, 31.6.

^{31}P NMR (400 MHz, CD_3OD) δ 4.8, 1.7, -0.1, -0.2.

ESI-TOF m/z calculated for $C_{48}H_{49}ClF_3N_4O_{14}P$ (M - H) − 1185.0990, 1185. 0579 observed.

$\varepsilon_{275} = 1.96 \times 10^4\ M^{-1}cm^{-1}$, $\varepsilon_{370} = 3.81 \times 10^3\ M^{-1}cm^{-1}$



72

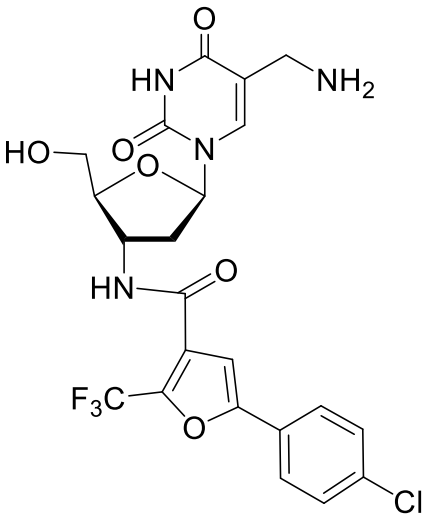
The starting material was stored as a 25 mM stock solution in DMF at -20 ºC. An aliquot of the starting material (2 µL, 50 nmol) was placed in a microtiter well plate and cooled to 4 ºC. Cold H_2O (3 µL) and NBS in MeCN (25 mM, 5 µL) was added and the well plate was shaken at 4 ºC for 9 min. After 9 min, the reaction was quenched by an equal volume of $Na_2S_2O_3$ (200 mM in H_2O) and well plate was shaken for additional 5-10 min. The reaction was concentrated by speed vacuum, redissolved in 1:1 MeCN/H_2O and used directly.

Reaction Conditions: [SM] = 5 mM, [NBS] = 12.5 mM, 30% H_2O, 20% DMF, 50% MeCN

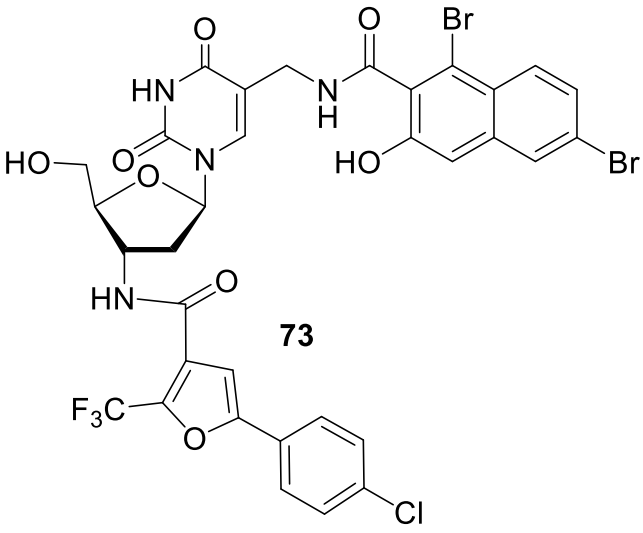
Quenching Conditions: [SM] = 2.5 mM, [NBS] = 6.25 mM, [$Na_2S_2O_3$] = 100 mM, 65% H_2O, 10% DMF, 25% MeCN.

ESI-TOF m/z calculated for $C_{48}H_{49}ClF_3N_4O_{14}P$ (M + H) – 1049.9738, 1049.9699 observed.

Preparation of 73



Compound **68** (50 mg, 0.08 mmol) was dissolved in concentrated aqueous ammonia (0.8 mL). The reaction was capped and stirred at 25 °C. After 5 h, TLC (5% MeOH in DCM, R_f = 0.2, UV active) confirmed the disappearance of starting material and formation of a new, more polar spot. The reaction was concentrated to yield 45 mg of crude amide.



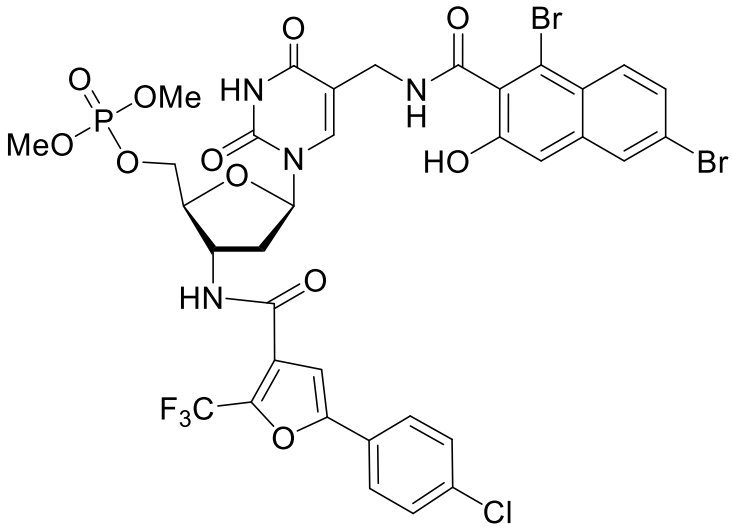
73

The crude amide (45 mg, 0.08 mmol) and NHS ester **X** (36 mg, 0.09 mmol. 1.1 eq) were azeotropically dried together in pyridine (2 x 0.5 mL). The contents were dissolved in DMF (1.6 mL). The reaction was stirred at 25 °C overnight. TLC (1:1 EtOAc/Hex, 2% MeOH, UV active, R_f = 0.2) suggested the starting materials were converted to product. The reaction was concentrated and purified by column chromatography (2:1 Hex/EtOAc -> 1:1 Hex/EtOAc -> 1:1 Hex/EtOAc -> 2% MeOH -> 1:1 Hex/EtOAc, 5% MeOH) to yield 53 mg (78%) of **73**. ^{1}H NMR (400 MHz, CD_3OD) δ 8.56 (s, 3H), 8.34 (s, 1H), 8.26 (s, 1H), 8.04 (m, 3H), 7.94 (d, *J* = 8 Hz, 1H), 7.86 (d, *J* = 8 Hz, 1H), 7.21 (s, 1H), 6.30 (t, *J* = 6.8 Hz, 1H), 4.69 (m, 1H), 4.34 (d, *J* = 3.2 Hz, 2H), 4.03 (d, *J* = 2.8 Hz, 1H), 3.87 (dd, *J* = 3.2, 13.9 Hz, 1H), 3.77 (dd, *J* = 2.8, 13.9 Hz, 1H), 2.45 (m, 2H).

^{13}C NMR (400 MHz, CD_3OD) δ 175.0, 127.9, 127.2, 118.28, 111.6, 55.7, 26.2, 19.2.

ESI-TOF m/z calculated for $C_{33}H_{24}Br_2ClF_3N_4O_8$ (M + H) – 854.9602, 854.9636 observed.

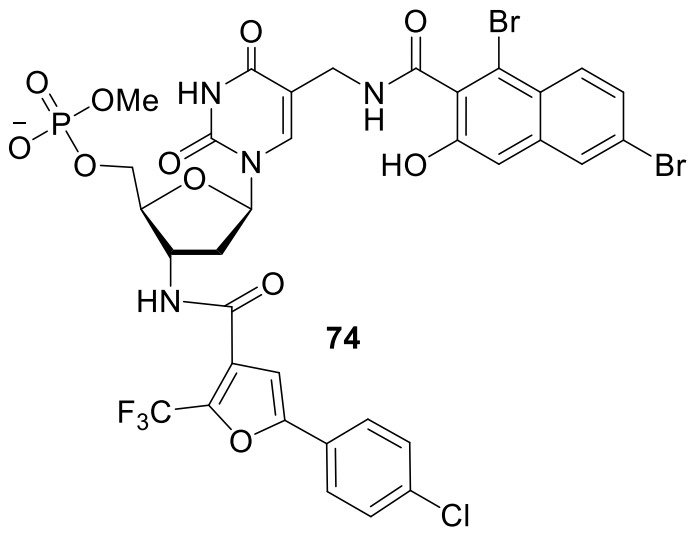
Preparation of 74



Compound **73** (20 mg, 0.023 mmol) and dimethyl N, N-diisopropyl phosphoramidite (6 mg, 6 µL, 0.03 mmol, 1.3 eq) were azeotropically dried together in pyridine (2 x 0.2 mL). S-ethyl tetrazole/MeCN (250 mM, 0.2 mL, 0.05 mmol, 2.2 eq) was added to the reaction flask. After 3.5 h, TLC (2:1 EtOAc/Hex) suggested the majority of **73** was consumed. tBuOOH (0.5 M, 7 mg, 15 µL, 0.07 mmol, 3 eq) was added for an additional 20 min. The reaction was concentrated and purified by column chromatography (1:1 Hex/EtOAc -> 1:1 Hex/EtOAc, 2% MeOH) to yield 8 mg (42%) of the intermediate phosphate triester. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.61 (s, 1H), 8.15 (d, *J* = 4.8 Hz, 1H), 7.96 (s, 1H), 7.87 (m, 2H), 7.53 (m, 4H), 7.33 (t, *J* = 4.8 Hz, 4H), 7.04 (s, 1H), 6.28 (m, 1H), 4.67 (m, 1H), 4.38 (m, 3H), 4.28 (m, 3H), 3.67 (s, 2H), 2.42 (m, 2H).

^{31}P NMR (400 MHz, $CDCl_3$) δ 10.6, 2.1, 0.0.

ESI-TOF m/z calculated for $C_{35}H_{29}Br_2ClF_3N_4O_{11}P$ (M + H) – 962.9578, 963.4038 observed.



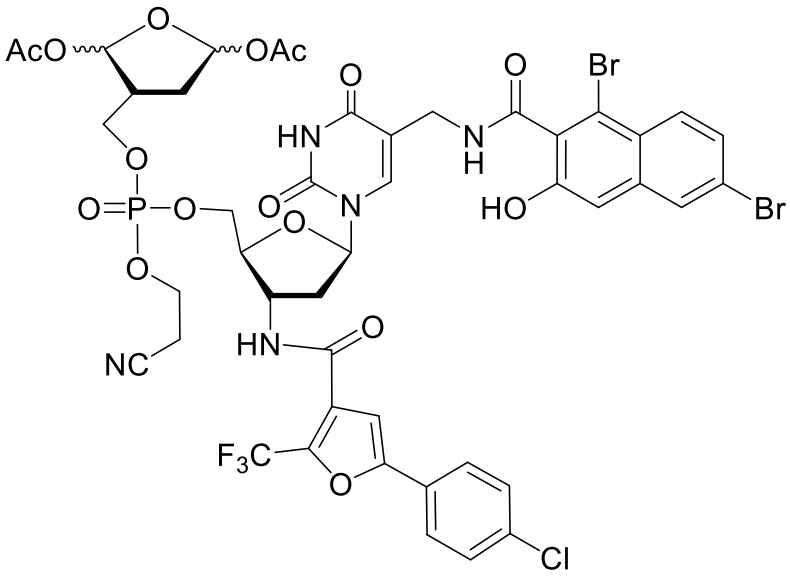
74

The intermediate phosphate triester (8 mg, 0.008 mmol) and NaI (2 mg, 0.008 mmol, 1 eq) were dissolved in acetone (0.1 mL). The reaction was stirred at 25 ^{o}C for 72 h. TLC (10% MeOH in DCM, R_f = 0.2) showed a new spot and loss of starting material. The reaction was concentrated and purified by column chromatography (EtOAc -> 25% Isopropanol in EtOAc -> 1:1 EtOAc/Isopropanol -> 2:2:1 EtOAc/ Isopropanol/H_2O) and Na^+ DOWEX column to yield 7.8 mg (100%) of **74**. ^{1}H NMR (400 MHz, CD_3OD) δ 7.75 (d, *J* = 8.8 Hz, 1H), 7.67 (m, 2H), 7.63 (d, *J* = 1.2 Hz, 1H), 7.50 (d, *J* = 8.8 Hz, 2H), 7.43 (dd, *J* = 1.2, 8.4 Hz, 2H), 7.26 (m, 2H), 6.90 (d, *J* = 8.4 Hz, 1H), 6.35 (t, *J* = 6 Hz, 1H), 5.23 (m, 1H), 4.27 (m, 5H), 3.63 (m, 2H), 3.54 (m, 1H), 3.49 (m, 2H), 2.87 (m, 3H).

^{13}C NMR (400 MHz, CD_3OD) δ 131.3, 124.9, 124.7, 118.6, 112.4, 52.6, 28.1, 28.0, 15.4.

^{31}P NMR (400 MHz, CD_3OD) δ 0.6.

Preparation of pro-72

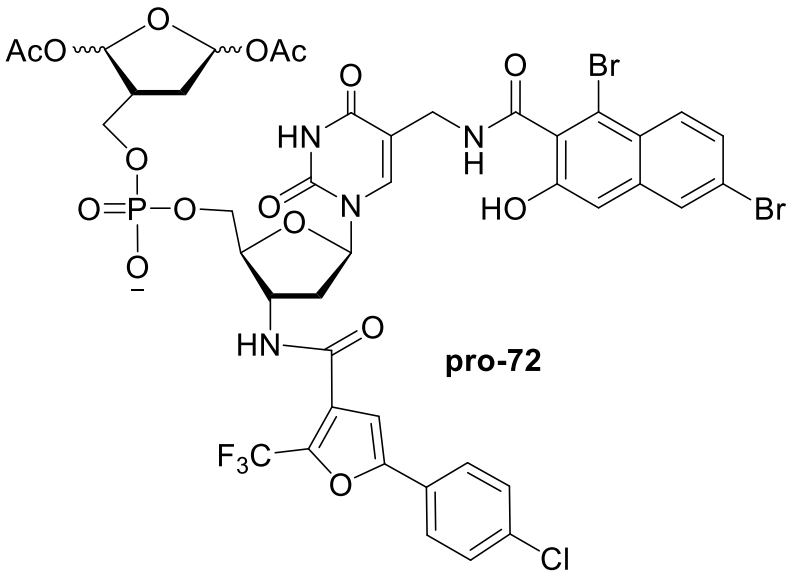


Compounds **73** (30 mg, 0.035 mmol) and **12** (35 mg, 0.042 mmol, 1.2 eq) were azeotropically dried together in pyridine (2 x 1 mL). Sethyl tetrazole/MeCN (250 mM, 0.25 mL, 0.053 mmol, 1.5 eq) was added to the reaction flask. After 4 h, TLC (1:1 EtOAc/Hex) suggested the majority of **73** was consumed. tBuOOH (0.5 M, 10 mg, 25 μL, 0.11 mmol, 3 eq) was added for 20 min. The reaction was concentrated and purified by column chromatography (1:1 Hex/EtOAc -> 1:1 Hex/EtOAc, 5% MeOH) to yield 25 mg (60%) of the intermediate phosphate triester. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.16 (m, 1H), 7.87 (s, 1H), 7.78 (d, *J* = 8.8 Hz, 1H), 7.50 (s, 1H), 7.29 (d, *J* = 8.8 Hz, 1H), 6.95 (s, 1H), 6.75 (m, 1H), 6.36 (quint, *J* = 3.2 Hz, 1H), 6.20 (s, 1H), 6.15 (m, 1H), 4.74 (q, *J* = 5.6 Hz, 1H), 4.31 (m, 4H), 4.23 (m, 6H), 3.82 (d, *J* = 5.6 Hz, 1H), 3.51 (m, 1H), 2.77 (m, 4H), 2.53 (m, 1H), 2.08 (s, 3H), 2.03 (s, 3H), 2.00 (m, 1H), 1.87 (m, 1H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 169.8, 100.8, 98.2, 60.2, 50.5, 32.2, 25.2, 20.9, 20.83, 20.77, 19.8, 19.7, 14.6, 14.0.

^{31}P NMR (400 MHz, $CDCl_3$) δ 17.3, 14.3, 9.2, 9.1, 8.1, 7.9, 7.8, 7.75, 7.70, 7.6, 7.5, -0.5, -0.9, -1.4, -1.9, -2.0, -2.5, -2.6, -2.7, -2.8.

ESI-TOF m/z calculated for $C_{45}H_{40}Br_2ClF_3N_5O_{16}P$ (M + H) – 1188.0215, 1188.0100 observed.



pro-72

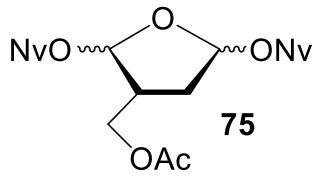
The intermediate phosphate triester (25 mg, 0.023 mmol) was dissolved in DMF (0.5 mL) and DIPEA (0.5 mL). The reaction was stirred at 25 °C overnight. The following morning, TLC (5% MeOH in DCM, R_f = 0.2, UV active) suggested product formation. The reaction was concentrated and purified by column chromatography (2% MeOH in DCM -> 5% MeOH in DCM) and Na^+ DOWEX to yield 12.4 mg (50%) of **pro-72**. ^{1}H NMR (800 MHz, CD_3OD) δ 8.53 (s, 1H), 8.12 (s, 1H), 8.02 (s, 1H), 7.92 (d, *J* = 8.8 Hz, 1H), 7.78 (d, *J* = 8.5 Hz, 2H), 7.56 (d, *J* = 8.8 Hz, 1H), 7.53 (d, *J* = 8.5 Hz, 2H), 7.24 (s, 1H), 6.30 (quint, *J* = 5.4 Hz, 1H), 6.25 (m, 1H), 4.73 (t, *J* = 6 Hz, 1H), 4.61 (q, *J* = 6 Hz, 1H), 4.49 (d, *J* = 14 Hz, 1H), 4.42 (dd, *J* = 3.6, 14 Hz, 1H), 4.24 (m, 3H), 4.14 (m, 1H), 4.05 (t, *J* = 7 Hz, 1H), 2.62 (m, 2H), 2.48 (q, *J* = 6 Hz, 2 H), 2.05 (m, 3H), 1.99 (s, 3H), 1.92 (s, 1H), 1.84 (m, 1H).

^{13}C NMR (800 MHz, $CDCl_3$) δ 175.7, 170.4, 170.1, 168.6, 163.8, 161.7, 157.5, 154.3, 150.7, 139.5, 138.6, 135.0, 133.9, 131.3, 130.8, 129.0, 128.4, 127.1, 126.5, 125.7, 125.3, 119.6, 118.3, 115.0, 111.1, 106.2, 101.1, 100.0, 98.6, 96.7, 85.7, 83.3, 65.0, 63.4, 59.7, 50.2, 43.2, 37.1, 35.4, 32.0, 26.7, 24.7, 22.6, 19.7, 19.6, 19.5, 14.2.

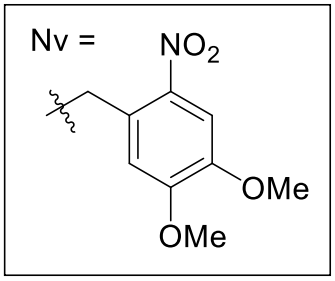
^{31}P NMR (400 MHz, CD_3OD) δ 6.4, 4.4, 4.3, 4.0, 3.9, 1.6, 1.5, 1.2, -0.26, -0.29, -0.4, -0.5.

ESI-TOF m/z calculated for $C_{42}H_{37}Br_2ClF_3N_4O_{16}P$ (M - H) – 1132.9950, 1133.0017 observed.

Preparation of 75

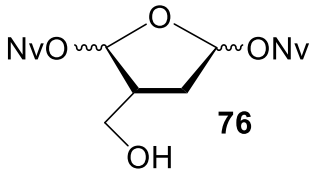


75



BF_3•etherate was distilled from CaH_2 at vacuum pressure. Compound **7** (1.73 g, 6.6 mmol, 1 eq) was dissolved in dry DCM (10 mL) and cooled to 0 °C. 6-Nitro veratryl alcohol (NvOH) (8.5 g, 40 mmol, 8 eq) was added to the reaction. BF_3•etherate (5.7 g, 5 mL, 40 mmol, 8 eq.) was added slowly over a period of 15-20 min. After 30 min, TLC (1:1 Hex/EtOAc, $R_f = 0.3$) suggested reaction was complete. The reaction was quenched with sat. $NaHCO_3$, diluted with DCM, and washed with sat. $NaHCO_3$ (1x10 mL). The aqueous layer was extracted with DCM (4x 15 mL) and the organic layer was washed with water (1x 30 mL), brine (1x30 mL) and dried over Na_2SO_4. The residue was concentrated and purified by column chromatography (5% EtOAc in DCM) to give 1.8 g (48%) of **75**. This product was difficult to purify away from the NvOH reagent so the impure product was carried forward to the next step. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.69 (s, 3H), 7.17 (s, 2H), 5.2 (m, 2H), 4.94 (dd, J = 0.5, 6.6 Hz, 3H), 3.99 (s, 3H), 3.97 (m, 5H), 3.94 (s, 4H), 3.91 (s, 3H), 3.83 (d, J = 6.6 Hz, 1H), 2.63 (t, J = 6.6 Hz, 1H), 2.04 (m, 3H).

Preparation of 76

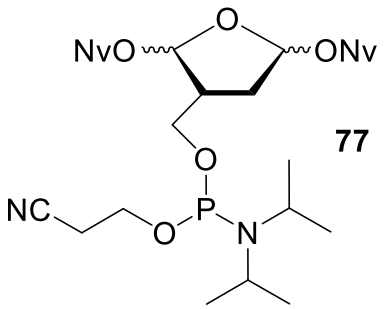


76

Compound **75** (1.8 g, 3.2 mmol) was dissolved in MeOH (3 mL). Sodium methoxide stock solution (2.5 M) was prepared by dissolving Na metal (580 mg, 25.2 mmol) in MeOH (10 mL). An aliquot of 2.5 M NaOMe (6.5 mL) was added to the reaction solution slowly (effectively diluting the NaOMe to 1.7 M). After 2 h, TLC (25% EtOAc in DCM, $R_f = 0.3$) confirmed reaction complete. The reaction was quenched with AcOH until neutral pH, diluted with DCM (20 mL), and washed with H_2O (2 x15 mL). The aqueous layer was extracted with DCM (4 x 20 mL) and the organic layer was washed with brine (1 x 30 mL) and dried over Na_2SO_4. The reaction was concentrated and purified by column chromatography (DCM

-> 25% EtOAc in DCM) to yield 1.1 g (70%) of **76**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.63 (s, 2H), 7.12 (m, 2H), 5.05 (m, 2H), 4.91 (m, 3H), 3.93 (s, 6H), 3.88 (s, 6H), 3.80 (m, 2H), 2.25 (m, 2H), 2.41 (m, 1H), 2.24 (m, 1H).

Preparation of 77

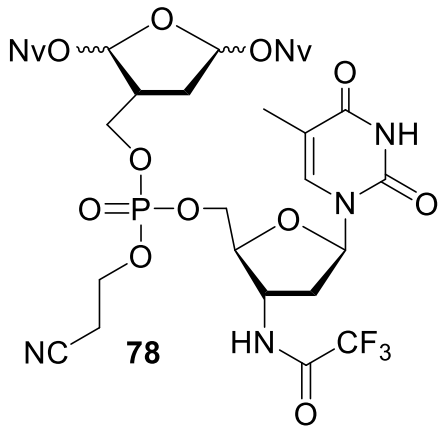


77

Compound **76** (279 mg, 0.53 mmol) was azeotropically dried with pyridine (2 x 2 mL). The flask cooled to 0 °C. DIPEA (284 mg, 0.4 mL, 2.2 mmol, 4 eq) was added to the cold starting material and the reactants dissolved in DCM (2 mL, 250 mM). 2-Cyanoethyl-N, N-diisopropylchlorophosphoramidite (150 mg, 0.15 mL, 0.6 mmol. 1.2 eq) was added. After 2 h, TLC (3:1 Hex/EtOAc, R_f = 0.3, stained with PAA) showed complete conversion to the phosphoramidite. The reaction was diluted with EtOAc (10 mL). The organic layer was washed with saturated bicarbonate solution (2 x15 mL) and the aqueous layers were extracted with EtOAc (2 x 20 mL). The combined organic layers were washed with brine (1 x 25 mL) and dried over Na_2SO_4. The organic layer was concentrated under vacuum and purified by column chromatography (3:1 hexanes/EtOAc) yielding 312 mg (81%) of **77**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.72 (s, 2H), 7.64 (s, 1H), 7.23 (t, *J* = 6.4 Hz, 1H), 7.13 (s, 1H), 5.43 (m, 1H), 5.27 (m, 1H), 5.00 (m, 5H), 4.02 (m, 4H), 3.97 (s, 3H), 3.94 (s, 3H), 3.83 (d, *J* = 6.4 Hz, 2H), 3.61 (m, 2H), 2.66 (m, 3H), 2.19 (s, 1H), 2.06 (s, 1H), 1.21 (m, 12H).

^{31}P NMR (400 MHz, $CDCl_3$) δ 148.3, 148.22, 148.17, 148.0, 147.9.

Preparation of 78



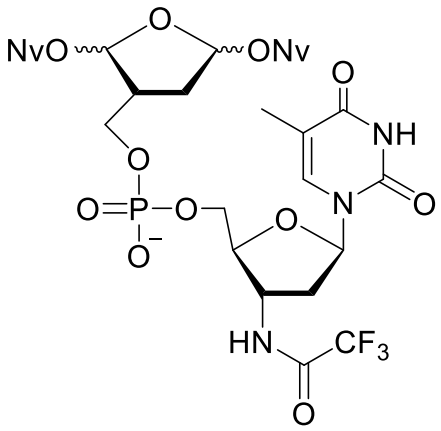
Compounds **77** (82 mg, 0.16 mmol) and **55** (182 mg, 0.34 mmol, 2.2 eq) were azeotropically dried together in pyridine (2 x 1 mL). S-ethyl tetrazole/MeCN (250 mM, 0.75 mL, 0.19 mmol, 1.2 eq) was added to the reaction flask. After 4 h, TLC (25% EtOAc in DCM, R_f = 0.5) suggested the majority of **55** was consumed. tBuOOH (0.5 M, 21 mg, 50 µL, 0.23 mmol, 1.5 eq) was added for 20 min. The reaction was evaporated to dryness under vacuum and purified by column chromatography (DCM -> 5% MeOH in DCM) to yield 77 mg (60%) of **78**. ^{1}H NMR (400 MHz, CDCl3) δ 10.2 (s, 1H), 8.8 (q, *J* = 6 Hz, 1H), 7.63 (s, 1H), 7.43 (s, 1H), 7.10 (m, 1H), 6.33 (s, 1H), 5.37 (m, 1H), 5.17 (m, 1H), 4.94 (m, 2H), 4.85 (m, 1H), 4.57 (s, 1H), 4.35 (s, 2H), 4.28 (m, 3H), 4.20 (s, 2H), 4.09 (m, 1H), 4.02 (m, 1H), 3.95 (s, 6H), 3.89 (s, 3H), 2.78 (q, *J* = 6 Hz, 2H), 2.65 (m, 1H), 2.51 (m, 1H), 2.27 (m, 2H), 1.87 (s, 3H), 1.79 (m, 1H), 1.21 (m, 2H).

^{13}C NMR (400 MHz, $CDCl_3$) δ 164.1, 155.2, 153.5, 153.4, 150.8, 149.0, 147.8, 124.1, 117.1, 108.0, 56.4, 56.3, 56.2, 53.4, 27.1, 19.7, 14.8, 12.3.

^{31}P NMR (400 MHz, $CDCl_3$) δ 14.0, 8.1, -2.0, -2.10, -2.11, -2.14, -2.2, -2.21, -2.3, -2.4, -2.5, -3.0, -3.6.

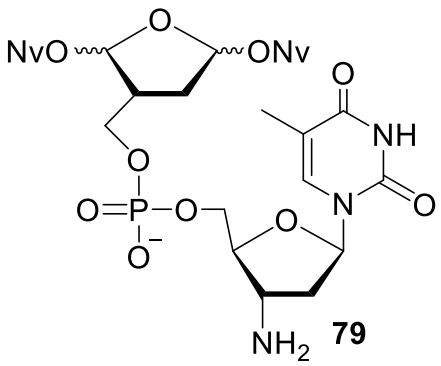
ESI-TOF m/z calculated for $C_{38}H_{44}F_3N_6O_{19}P$ (M + H) – 977.2351, 977.2365 observed.

Preparation of 79



Compound **78** (77 mg, 0.08 mmol) was dissolved in MeCN (0.5 mL) and TEA (0.5 mL). The reaction was stirred at 25 °C for 16 h. TLC (5% MeOH in DCM, R_f = 0.3) suggested reaction was complete. The reaction was concentrated to yield 70 mg of crude material. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.55 (s, 2H), 7.45 (s, 1H), 7.05 (s, 2H), 6.91 (m, 1H), 6.18 (s, 1H), 5.32 (s, 1H), 5.22 (s, 1H), 5.10 (s, 1H), 4.87 (m, 1H), 4.76 (m, 2H), 4.50 (m, 1H), 4.14 (s, 2H), 3.97 (m, 3H), 3.87 (s, 4H), 3.82 (s, 4H), 3.79 (t, *J* = 4 Hz, 3H), 3.71 (s, 2H), 2.54 (m, 1H), 2.40 (m, 3H), 1.76 (s, 3H).

ESI-TOF m/z calculated for $C_{35}H_{40}F_3N_5O_{19}P$ (M - H) – 923.2013, 923.2099 observed.



79

The crude material (32 mg, 0.035 mmol) was further deprotected by addition of NaOH (1.7 mL, 0.1 M, 5 eq.) in 20% MeCN in H_2O. After 2 h, the reaction was quenched by an equal molar equivalent of formic acid (17 μL, 1 M). The reaction was evaporated to dryness to yield 29 mg (100%) of **79**. ^{1}H NMR (400 MHz, CD_3OD) δ 7.78 (m, 1H), 7.76 (m, 1H), 7.67 (m, 2H), 7.55 (m, 1H), 7.22 (m, 2H), 6.20 (m, 1H), 5.40 (m, 1H), 5.29 (m, 1H), 5.17 (m, 1H), 5.00 (m, 1H), 4.82 (m, 2H), 4.07 (s, 3H), 3.94 (m, 6H) 3.89 (s, 5H), 3.85 (s, 2H), 3.77 (m, 2H), 3.59 (m, 1H), 2.86 (s, 1H), 2.26 (m, 2H), 2.17 (m, 2H), 1.84 (m, 3H).

^{31}P NMR (400 MHz, CD_3OD) δ 4.7, 0.44, 0.41, 0.36, 0.32, 0.27, 0.22, 0.17, 0.12, 0.07.

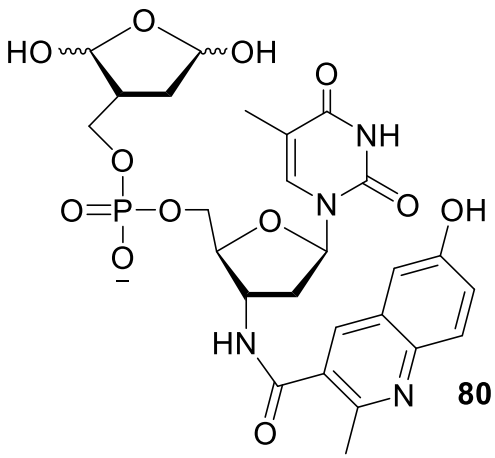
ESI-TOF m/z calculated for $C_{33}H_{41}N_5O_{18}P$ (M + H) – 829.2262, 829.2459 observed.

General Procedure for preparing Library 58 through Scaffold 79

Amine scaffold **79** (50 nmol) was azeotropically dried with carboxylic acid (140 nmol, 1.4 eq) in pyridine (1 x 15 μL) using a Speed Vac concentrator in a 384-well microtiter plate (VWR). To each well, a 2X activating solution (5 μL; 28 mM HBTU and 28 mM HOBt in DMF), DIPEA (2 μL), and DMF (3 μL) were added. The final concentrations during reaction were: [**79**] = 5 mM, [acid] = 7 mM, [HBTU] = 7 mM, [HOBt] = 7 mM, 20% DIPEA in DMF. The well plate was shaken at 25 ºC overnight. Some wells were analyzed by ESI-MS to confirm coupling efficiency. The solutions were concentrated using a Speed Vac concentrator and the well plate was covered and stored at -80 ºC.

Immediately before an assay, the amide was thawed, dissolved in 1:1 MeCN/H_2O (10 μL, 5 mM). Samples were photolyzed in a clear 384-well plate (VWR) using a transilluminator at 365 nm for 15 min. During photolysis, the plate was covered with aluminum foil and a fan was turned on near the system to prevent heating the samples. Prior to carrying out the photolysis on the entire library, the photolysis reaction was carried out on model compound (**79**) for 10-30 min and analyzed by ESI and HPLC, to determine optimal time. The reaction was complete after 15 min. There was no significant difference in reaction products after longer photolysis times (i.e. 30 min). It should be noted that prior to use in an assay, the sample was diluted to a final volume of 10 μL to account for solvent evaporation during photolysis. Some samples were analyzed ESI-MS to confirm product formation.

Preparation of 80



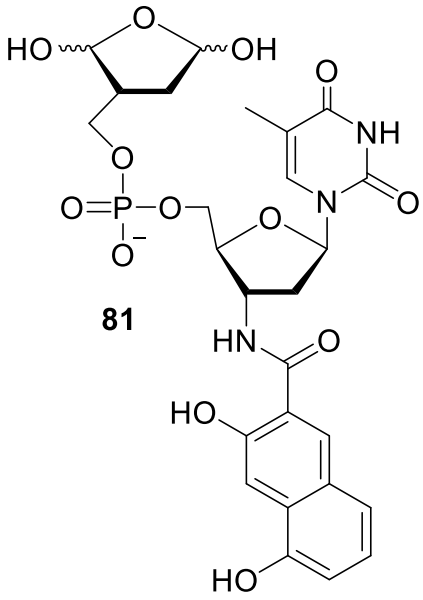
80

Scaffold **79** (10 mg, 0.01 mmol) and 6-hydroxy-2-methylquinoline-3-carboxylic acid (6 mg, 0.022 mmol, 2 eq) were azeotropically dried together in pyridine (2 x 0.5 mL). HBTU (6 mg, 0.02 mmol, 1.4 eq) and HOBT (2 mg, 0.02 mmol, 1.4 eq) were added to the reaction flask. The contents were dissolved in DMF (0.2 mL) and DIPEA (40 μL). The reaction was stirred at 25 ºC overnight. TLC (5% MeOH in DCM, R_f = 0.3) suggested the starting materials were converted to product. The reaction was evaporated to dryness under vacuum and purified by column chromatography (DCM -> 5% MeOH in DCM) to yield 10 mg (90%) of **84**.

A small portion of **84** (50 nmol, 5 mM in 1:1 MeCN/H_2O) was photolyzed at 365 nm for 15 min. The product was confirmed by analyzing an aliquot of the reaction (2 μL) by ESI-MS.

ESI-MS m/z calculated for $C_{26}H_{31}N_4O_{12}P$ (M - H) – 621.17, 621.25 observed.

Preparation of 81



81

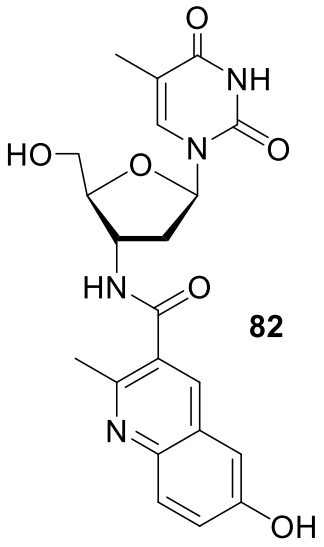
Scaffold **79** (10 mg, 0.01 mmol) and **52** (5.8 mg, 0.014 mmol, 1.4 eq) were azeotropically dried together with pyridine (2 x 0.5 mL). The contents were dissolved in DMF (0.2 mL) and DIPEA (40 μL). The reaction was stirred at 25 ºC overnight. TLC (5% MeOH in DCM, R_f = 0.3) suggested the starting materials were converted to product. The reaction was evaporated to dryness under vacuum and purified by column

chromatography (DCM -> 5% MeOH in DCM) to yield 6 mg (60%) of **89**.

A small portion of **89** (50 nmol, 5 mM in 1:1 MeCN/H_2O) was photolyzed at 365 nm for 15 min. The product was confirmed by analyzing an aliquot of the reaction (2 μL) by ESI-MS.

ESI-MS m/z calculated for $C_{26}H_{30}N_3O_{13}P$ (M - H) – 622.15, 622.10 observed.

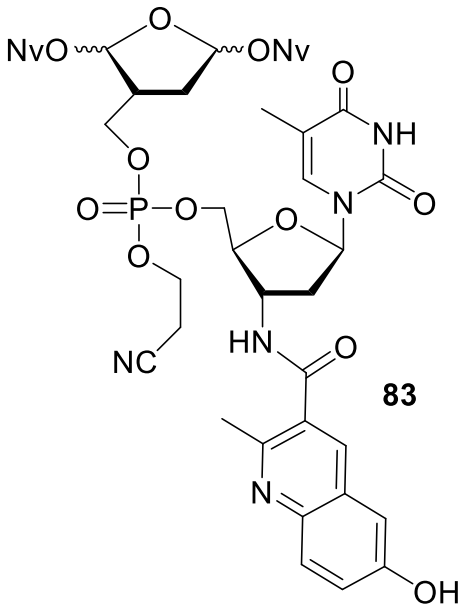
Preparation of 82



82

The precursor to **55** (3'-amino deoxythymidine) (210 mg, 0.87 mmol, 1.6 eq) and 6-hydroxy-2-methylquinoline-3-carboxylic acid (125 mg, 0.52 mmol) were azeotropically dried together with pyridine (1 x 1 mL). HBTU (280 mg, 0.73 mmol, 1.4 eq) and HOBt (100 mg, 0.73 mmol, 1.4 eq) were added and reagents were dissolved in DMF (6 mL) and DIPEA (1.5 mL, 20%). Reaction was stirred at 25 ºC overnight. The following day, TLC (3% MeOH, R_f = 0.7) suggested the carboxylic acid was consumed. The reaction was concentrated and purified by column chromatography (DCM -> 3% MeOH in DCM) to yield 156 mg (65%) **82**. ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.3 (s, 1H), 10.00 (s, 1H), 8.95 (d, J = 7.4 Hz, 1H), 8.15 (s, 1H), 7.83 (s, 1H), 7.79 (d, J = 9 Hz, 1H), 7.33 (d, J = 9 Hz, 1H), 7.17 (s, 1H), 6.28 (t, J = 6.5 Hz, 1H), 5.17 (s, 1H), 4.58 (s, 2H), 4.09 (s, 2H), 3.94 (s, 1H), 3.75 (m, 2H), 2.29 (m, 2H), 1.81 (s, 3H).

Preparation of 83

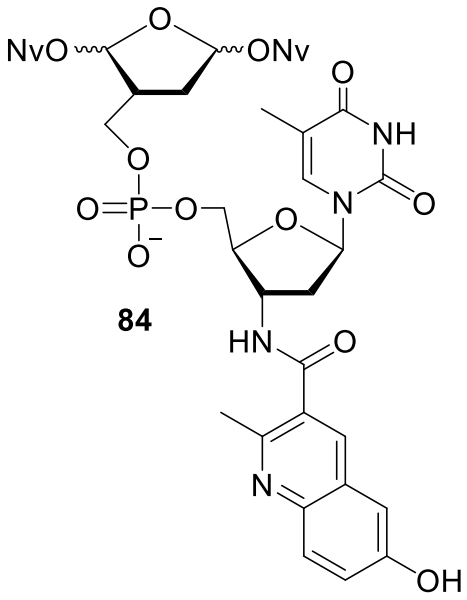


83

Compounds **77** (63 mg, 0.09 mmol) and **82** (92 mg, 0.22 mmol, 2.5 eq) were azeotropically dried together in pyridine (2 x 1 mL). S-ethyl tetrazole/MeCN (250 mM, 1 mL, 0.25 mmol, 2.7 eq) and DCM (1 mL) were added to the reaction flask. After 4 h, TLC (3% MeOH in DCM, R_f = 0.5) suggested the majority of **82** was consumed. tBuOOH (0.5 M, 24 mg, 50 µL, 0.26 mmol, 2.8 eq) was added for 20 min. The reaction was evaporated to dryness under vacuum and purified by column chromatography (DCM -> 3% MeOH in DCM) to yield 50 mg (53%) of **83**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.78 (s, 1H), 8.17 (s, 1H), 7.73 (d, *J* = 7.5 Hz, 1H), 7.64 (d, *J* = 7.5 Hz, 1H), 7.54 (m, 1H), 7.29 (m, 2H), 6.99 (m, 2H), 6.31 (s, 1H), 5.33 (s, 2H), 5.04 (m, 1H), 4.87 (m, 2H), 4.72 (m, 1H), 4.47 (s, 2H), 4.31 (s, 3H), 3.92 (s, 3H), 3.88 (m, 5H), 3.69 (m, 3H), 3.17 (q, *J* = 6.6 Hz, 2H), 3.14 (q, *J* = 7.4 Hz, 2H), 2.79 (s, 2H), 2.70 (s, 3H), 2.57 (s, 1H), 2.36 (s, 1H), 2.16 (s, 1H), 1.77 (s, 3H),1.32 (t, *J* = 6.6, 4H).

^{31}P NMR (400 MHz, $CDCl_3$) δ -2.0, -2.2, -2.3.

Preparation of 84



84

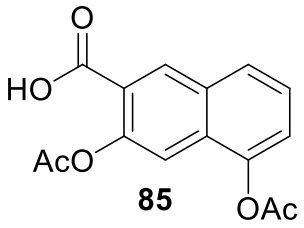
Compound **83** (20 mg, 0.02 mmol) was dissolved in 50% TEA in MeCN (0.5 mL). The reaction was stirred overnight at 25 °C. The following day, the reaction was concentrated and purified by a DOWEX Na^+ exchange column to yield 18 mg (94%) of **84**. ^{1}H NMR (800 MHz, CD_3OD) δ 8.17 (m, 1H), 7.80 (t, *J* = 8 Hz, 1H), 7.69 (dd, *J* = 7.2, 8 Hz, 2H), 7.52 (m, 1H), 7.35 (m, 1H), 7.27 (quint, *J* = 3 Hz, 2H), 7.16 (quint, *J* = 3 Hz, 1H), 7.07 (q, *J* = 7.2 Hz, 1H), 6.35 (m, 1H), 6.25 (m, 1H), 5.42 (m, 3H), 5.10 (m, 2H), 4.88 (m, 2H), 4.70 (m, 2H), 4,65 (quint, *J* = 5 Hz, 1H), 4.25 (t, *J* = 4 Hz, 1H), 4.22 (q, *J* = 4 Hz, 1H), 4.19 (m, 2H), 4.14 (q, *J* = 5 Hz, 1H), 4.05 (m, 1H), 3.97 (d, *J* = 5 Hz, 3H), 3.87 (m, 4H), 3.77 (m, 2H), 2.70 (m, 3H), 2.52 (m, 1H), 2.44 (m, 1H), 2.23 (m, 2H), 2.11 (m, 1H), 1.98 (s, 1H), 1.93 (m, 2H). ^{13}C NMR (800 MHz, CD_3OD) δ 169.7, 165.0, 164.9, 157.5, 155.9, 152.5, 147.7, 143.1, 142.3, 139.1, 134.2, 130.0, 127.5, 127.3, 123.7, 123.5, 122.9, 117.4, 110.6, 110.2, 110.0, 107.6, 107.4, 106.8, 104.9, 104.8, 104.0, 84.6, 83.4, 66.8, 66.4, 65.9, 65.5, 65.2, 55.5, 55.3, 55.2, 50.7, 41.2, 39.9, 39.0, 37.1, 33.9, 32.8, 26.7, 21.4, 14.1.

^{31}P NMR (400 MHz, CD_3OD) δ -2.0, -2.1, -2.3.

ESI-TOF m/z calculated for $C_{44}H_{49}N_6O_{20}P$ (M + H) – 1013.2739, 1013.2813 observed.

$\varepsilon_{316} = 5.55 \times 10^4\ M^{-1}cm^{-1}$

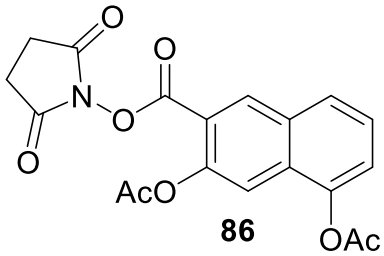
Preparation of 85[199]



85

3, 5-Dihydroxy naphthoic-2-acid (122 mg, 0.6 mmol, 1 eq) and DMAP (7.17 mg, 0.06 mmol, 0.1 eq) were dissolved in DCM (1.5 mL) and cooled to 0 ºC. Pyridine (559 mg, 0.57 mL, 7.03 mmol, 12.5 eq) was added slowly to the solution. Next, acetic anhydride (358 mg, 0.33 mL, 3.5 mmol, 6 eq) was added dropwise. After 2.5 h, TLC (10% MeOH in DCM, R_f = 0.7) confirmed the reaction was complete. The reaction was quenched with sat. $NaHCO_3$ until neutral pH. The mixture was diluted with EtOAc (5 mL) and washed with water (2 x 10 mL). The aqueous layer was extracted with EtOAc (2 x 20 mL), the combined organic layers were washed with brine (1 x 30 mL) and dried over Na_2SO_4. The crude residue was purified by column chromatography (1:1 Hex/EtOAc-> 1:2 Hex/EtOAc) to yield 40 mg (20%) of **85.** ^{1}H NMR (400 MHz, CD_3CN) δ 8.68 (s, 1H), 7.94 (d, *J* = 8 Hz, 1H), 7.65 (s, 1H), 6.58 (t, *J* = 8 Hz, 1H), 7.41 (d, *J* = 8 Hz, 1H), 2.43 (s, 3H), 2.32 (s, 3H).

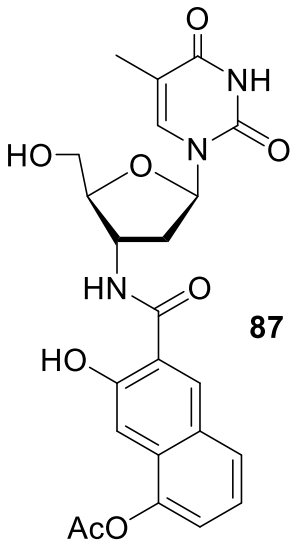
Preparation of 86



86

Compound **85** (630 mg, 2.2 mmol) was azeotropically dried with pyridine (2 x 1 mL). N-hydroxysuccinimide (360 mg, 3.3 mmol, 1.5 eq) and EDC (640 mg, 3.3 mmol, 1.5 eq) were added and the reagents were dissolved in DMF (4 mL). The reaction stirred at 25 ºC overnight. The following day, TLC (3% MeOH in DCM, R_f = 0.8) suggested the starting material was consumed. The reaction was concentrated and purified by column chromatography (3:1 Hex/EtOAc -> 1:1 Hex/EtOAc) to yield 503 mg (60%) of **86**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 876 (s, 1H), 7.83 (d, *J* = 8.2 Hz, 1H), 7.65 (s, 1H), 7.55 (dd, *J* = 7.7, 8.2 Hz, 1H), 7.46 (d, *J* = 7.7 Hz, 1H), 2.89 (s, 4H), 2.44 (s, 3H), 2.35 (s, 3H).

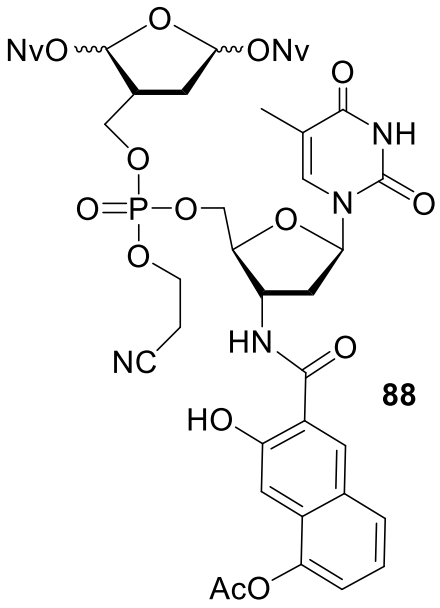
Preparation of 87



87

The precursor to **55** (3'-amino deoxythymidine) (260 mg, 1.1 mmol, 2.4 eq) and **86** (176 mg, 0.46 mmol) were azeotropically dried together with pyridine (1 x 1 mL). The reagents were dissolved in DMF (2.2 mL) and DIPEA (0.6 mL, 20%). The reaction was stirred at 25 ºC overnight. The following day, TLC (3% MeOH, $R_f = 0.7$) suggested the carboxylic acid was consumed. The reaction was concentrated and purified by column chromatography (DCM -> 3% MeOH in DCM) to yield 120 mg (45%) **87**. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.05 (d, J = 6.8 Hz, 1H), 8.00 (s, 1H), 7.90 (d, J = 9 Hz, 1H), 7.38 (d, J = 2.2 Hz, 1H), 7.35 (dd, J = 2.2, 9 Hz, 1H), 7.22 (s, 1H), 6.12 (t, J = 6.8 Hz, 1H), 4.56 (m, 1H), 4.35 (s, 2H), 4.14 (m, 1H), 3.37 (s, 1H), 2.67 (s, 2H), 2.27 (s, 2H), 2.12 (m, 1H), 2.06 (s, 3H), 1.59 (s, 3H).

Preparation of 88



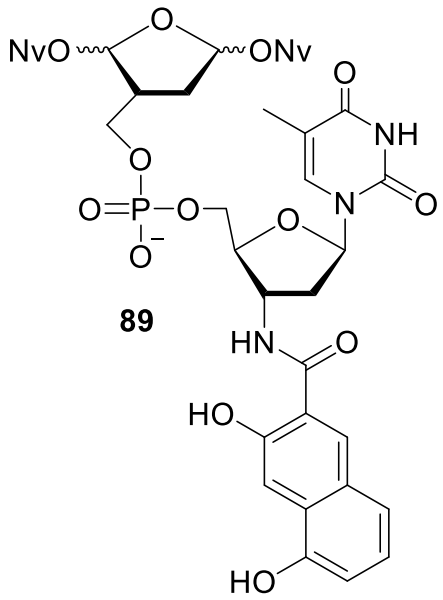
88

Compounds **77** (30 mg, 0.04 mmol) and **87** (35 mg, 0.07 mmol, 1.75 eq) were azeotropically dried together in pyridine (2 x 0.5 mL). S-ethyl tetrazole/MeCN (250 mM, 0.4 mL, 0.1 mmol, 2.5 eq) was added to the reaction flask. After 4 h, TLC (3% MeOH in DCM, $R_f = 0.5$) suggested the majority of **87** was consumed. tBuOOH (0.5 M, 12 mg, 25 µL, 0.13 mmol, 3.2 eq) was added for 20 min. The reaction was evaporated to dryness under vacuum and purified by column chromatography (DCM -> 3% MeOH in DCM) to yield 25 mg (60%) of **88**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.30 (d, J = 5 Hz, 2H), 7.60 (m, 2H), 7.46 (m, 1H),

7.20 (m, 2H), 7.05 (m, 1H), 6.27 (d, J = 5 Hz, 1H), 5.37 (m, 2H), 5.06 (m, 1H), 4.91 (m, 3H), 4.73 (m, 2H), 4.46 (m, 3H), 4.29 (m, 4H), 3.95 (m, 4H), 3.88 (m, 5H), 3.82 (m, 2H), 2.80 (m, 2H), 2.65 (s, 6H), 2.54 (m, 1H), 2.43 (s, 3H), 2.17 (s, 1H), 1.87 (m, 2H).

^{31}P NMR (400 MHz, $CDCl_3$) δ -2.06, -2.11, -2.16, -2.19, -2.23, -2.26, -2.31.

Preparation of 89



89

Compound **88** (20 mg, 0.02 mmol) was dissolved in 50% TEA in MeCN (0.5 mL). The reaction was stirred overnight at 25 °C. The following day, the reaction was concentrated and purified by a DOWEX Na^+ exchange column to yield 19 mg (95%) of **89**. ^{1}H NMR (800 MHz, CD_3OD) δ 8.31 (s, 1H), 7.84 (s, 1H), 7.59 (m, 2H), 7.46 (m, 2H), 7.28 (s, 2H), 7.09 (m, 2H), 7.01 (m, 1H), 6.78 (m, 1H), 6.38 (m, 1H), 5.40 (m, 2H), 5.06 (m, 1H), 4.78 (m, 2H), 4.28 (m, 2H), 4.18 (m, 2H), 4.02 (m, 1H), 3.93 (m, 4H), 3.83 (m, 8H), 3.75 (m, 2H), 2.63 (s, 1H), 2.51 (m, 2H), 2.23 (m, 1H), 2.07 (m, 1H), 1.95 (m, 3H).

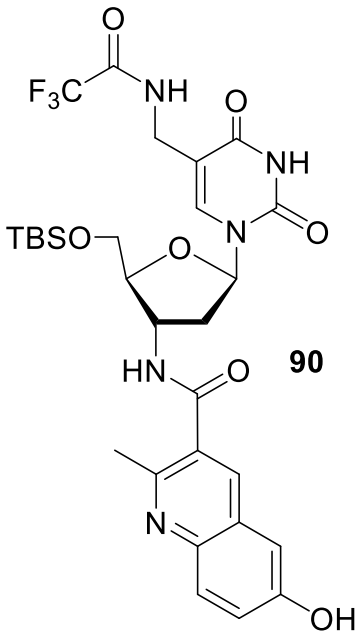
^{13}C NMR (800 MHz, CD_3OD) δ 169.3, 165.0, 157.5, 154.4, 153.3, 151.5, 150.9, 147.6, 139.9, 136.6, 129.75, 129.68, 129.1, 128.4, 128.3, 110.6, 110.0, 109.9, 109.2, 107.7, 107.4, 105.8, 104.8, 84.6, 83.9, 66.3, 66.2, 65.2, 55.5, 55.4, 55.2, 39.0, 37.3, 33.9, 29.4, 26.8, 24.8, 14.2, 11.3.

^{31}P NMR (400 MHz, CD_3OD) δ -2.0, -2.1, -2.16, -2.19, -2.23, -2.26, -2.30.

ESI-TOF m/z calculated for $C_{44}H_{49}N_5O_{21}P$ (M + H) – 1014.2579, 1014.2628 observed.

$\varepsilon_{348} = 6.98 \times 10^3\ M^{-1}cm^{-1}$

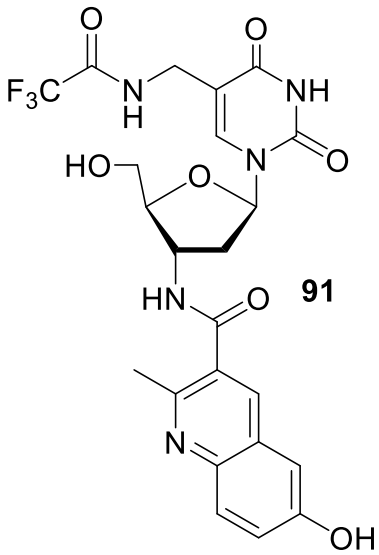
Preparation of 90



90

Compound **66** (340 mg, 0.73 mmol) was azeotropically dried with 6-hydroxy-2-methylquinoline-3-carboxylic acid (240 mg, 1.0 mmol, 1.4 eq) in pyridine (2 x 1 mL). HBTU (370 mg, 1.0 mmol, 1.4 eq) and HOBt (150 mg, 1.0 mmol, 1.4 eq) were added. Reagents were dissolved in DMF (3.2 mL, 200 mM) with 20% DIPEA (0.8 mL). The reaction was stirred at 25 °C overnight. The following day, the reaction was concentrated and purified by column chromatography to yield 328 mg (70%) of **90**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.07 (s, 1H), 7.79 (s, 2H), 7.57 (s, 1H), 7.43 (m, 2H), 6.17 (t, *J* = 7.2 Hz, 1H), 4.53 (m, 1H), 4.07 (m, 2H), 4.00 (m, 1H), 3.93 (d, *J* = 2.4 Hz, 2H), 2.39 (s, 3H), 2.31 (s, 3H), 2.25 (m, 2H), 0.89 (s, 9H), 0.11 (d, *J* = 1.8 Hz, 6H).

Preparation of 91

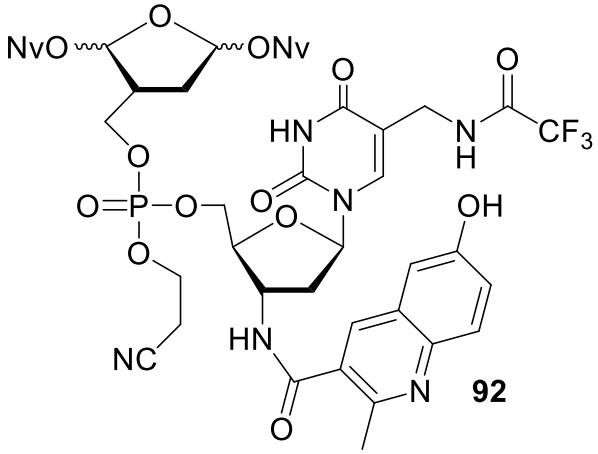


91

Compound **90** (160 mg, 0.25 mmol) was dissolved in THF (1.1 mL). TEA•3HF (99 mg, 0.62 mmol, 0.1 mL, 2.5 eq) was added to the flask. The reaction was stirred at 25 °C overnight. TLC (3% MeOH in DCM, R_f = 0.2, UV active and stained with PAA) confirmed conversion to product. The reaction was concentrated and purified by column chromatography (1:1 Hex/EtOAc -> 2% MeOH in 1:1 Hex/EtOAc) to yield 191 mg (75%) of **91**. ^{1}H NMR (400 MHz, CD_3OD) δ 8.2 (d, *J* = 10 Hz, 2H), 7.85 (d, *J* = 11.8 Hz, 2H), 7.47 (m, 1H), 7.38 (dd, *J* = 3.6, 11.8 Hz, 1H), 7.18 (d, *J* = 3.6 Hz, 1H), 6.27 (t, *J* = 8.3 Hz,

1H), 4.75 (q, *J* = 10 Hz, 1H), 4.18 (s, 2H), 3.97 (dd, *J* = 3.6, 16.3 Hz, 1H), 3.88 (dd, *J* = 5, 16.3 Hz, 1H), 3.72 (q, *J* = 10 Hz, 1H), 2.71 (s, 3H), 2.53 (t, *J* = 8.3 Hz, 2H).

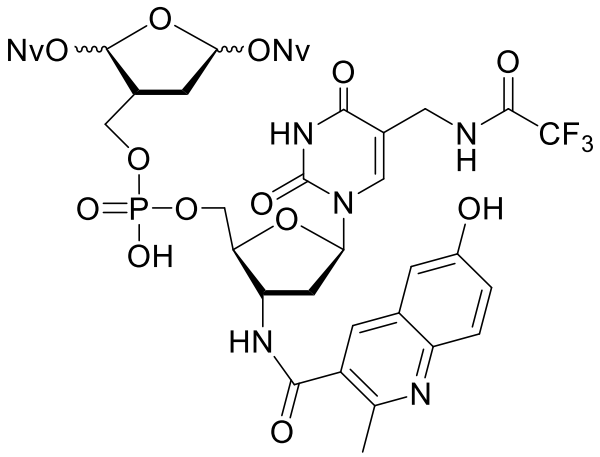
Preparation of 92



Compounds **77** (132 mg, 0.18 mmol, 4 eq) and **91** (25 mg, 0.04 mmol) were azeotropically dried together in pyridine (2 x 1 mL). S-ethyl tetrazole/MeCN (250 mM, 0.75 mL, 0.19 mmol, 4.5 eq) was added to the reaction flask. After 4 h, TLC (3% MeOH in DCM, R_f = 0.5) suggested the majority of **91** was consumed. tBuOOH (0.5 M, 21 mg, 50 μL, 0.2 mmol, 5 eq) was added for 20 min. The reaction was evaporated to dryness under vacuum and purified by column chromatography (DCM -> 5% MeOH in DCM) to yield 45 mg (95%) of **92**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.2 (m, 2H), 7.85 (m, 1H), 7.75 (m, 1H), 7.56 (m, 2H), 7.44 (m, 2H), 7.30 (m, 1H), 7.20 (m, 2H), 7.02 (m, 1H), 6.28 (s, 1H), 5.44 (m, 2H), 5.04 (m, 2H), 4.86 (m, 4H), 4.28 (m, 7H), 3.92 (s, 6H), 3.85 (m, 9H), 3.73 (s, 4H), 2.92 (m, 1H), 2.79 (s, 1H), 2.70 (s, 3H), 2.58 (m, 1H), 2.23 (m, 1H), 1.77 (m, 4H).

^{31}P NMR (400 MHz, $CDCl_3$) δ -2.0, -2.2, -2.3.

Preparation of 93

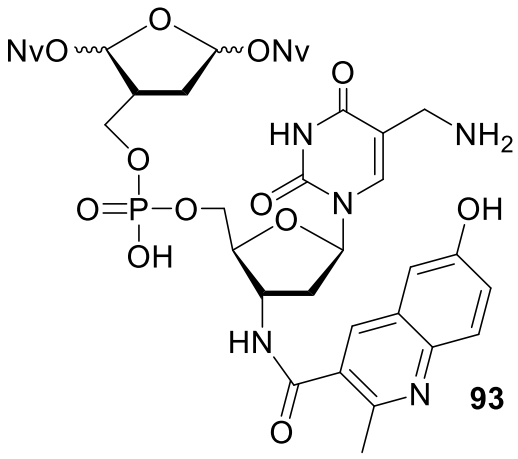


Compound **92** (45 mg, 0.04 mmol) was dissolved in MeCN (0.25 mL) and TEA (0.25 mL). The reaction stirred at 25 °C for 16 h. TLC (5% MeOH in DCM, $R_f = 0.3$) suggested reaction was complete. The reaction was concentrated to yield 30 mg (66%) crude material. The crude material was purified by DOWEX Na^+ ion exchange column and lyophilized three times to remove excess triethylamine. ^{1}H NMR (800 MHz, CD_3OD) δ 8.20 (m, 1H), 7.80 (m, 1H), 7.71 (m, 1H), 7.63 (d, *J* = 2.8 Hz, 1H), 7.54 (m, 2H), 7.34 (d, *J* = 2.8 Hz, 1H), 7.20 (s, 2H), 7.06 (m, 1H), 6.30 (s, 1H), 5.41 (m, 2H), 5.04 (m, 1H), 4.81 (m, 2H), 4.28 (m, 4H), 4.21 (m, 2H), 4.14 (m, 1H), 4.05 (q, *J* = 6 Hz, 1H), 3.97 (d, *J* = 14 Hz, 2H), 3.93 (d, *J* = 14 Hz, 3H), 3.87 (m, 4H), 3.87 (m, 4H), 3.76 (m, 2H), 2.80 (t, *J* = 6 Hz, 1H), 2.70 (d, *J* = 6 Hz, 2H), 2.67 (m, 1H), 2.52 (m, 2H), 2.11 (m, 1H), 1.88 (m, 1H), 1.31 (m, 3H).

^{13}C NMR (800 MHz, CD_3OD) δ 169.7, 163.5, 157.5, 157.4, 157.2, 156.3, 155.9, 153.7, 153.3, 152.5, 151.3, 150.7, 147.7, 142.3, 139.8, 139.7, 139.2, 139.1, 134.2, 130.0, 129.8, 128.2, 127.3, 126.2, 123.0, 116.9, 115.4, 110.1, 110.0, 109.5, 109.1, 108.5, 107.6, 107.4, 104.8, 85.4, 83.7, 66.3, 65.8, 64.8, 55.2, 37.6, 32.8, 29.4, 26.7, 23.4, 21.9, 21.8, 21.4, 19.0, 17.9, 14.2, 7.9.

^{31}P NMR (400 MHz, CD_3OD) δ -2.1, -2.16 -2.19, -2.23, -2.26 -2.30.

ESI-TOF m/z calculated for $C_{46}H_{49}F_3N_7O_{21}P$ (M + H) – 1124.2671, 1124.2737 observed.



The crude material (15 mg, 0.013 mmol) was dissolved in concentrated aqueous ammonium hydroxide (0.6 mL, 20 mM). After 2 h, the reaction was quenched by an equal molar equivalent of formic acid (12 μL, 1 M). The reaction was evaporated to dryness to yield 14 mg (100%) of **93**. The product was confirmed by UPLC-MS.

ESI-TOF m/z calculated for $C_{44}H_{50}N_7O_{20}P$ (M + H) – 1028.2848, 1028.2908 observed.

General Procedure for Amide Coupling and Deprotection for Library 94 (Includes the preparation of 95-102)

Amine scaffold **93** (50 nmol) was azeotropically dried with carboxylic acid (70 nmol, 1.4 eq) in pyridine (1 x 15 μL) using a Speed Vac concentrator in a 384-well microtiter plate (VWR). To each well, a 2X activating solution (5 μL; 14 mM HBTU and 14 mM HOBt in DMF), DIPEA (2 μL), and DMF (3 μL) were added. The final concentrations during reaction were: [**93**] = 5 mM, [acid] = 7 mM, [HBTU] = 7 mM, [HOBt] = 7 mM, 20% DIPEA in DMF. The well plate was shaken at 25 ºC overnight. Some wells were analyzed by ESI-MS to confirm coupling efficiency. The solutions were concentrated using a Speed Vac concentrator and the well plate was covered and stored at -80 ºC.

Immediately before an assay, the amide was thawed. The crude amide (50 nmol, 10 μL, 5 mM in 1:1 MeCN/H_2O) was photolyzed at 365 nm for 15 min using a transilluminator. It should be noted that prior to use in an assay, the sample was diluted to a final volume of 10 μL to account for solvent evaporation during photolysis. The product was confirmed by analyzing an aliquot of the reaction (2 μL) by ESI-MS.

References

(1) Salehan, M. R.; Morse, H. R. DNA Damage Repair and Tolerance: A Role in Chemotherapeutic Drug Resistance. *Br J Biomed Sci* **2013**, *70* (1), 31–40.

(2) Kelley, M.; Logsdon, D.; Fishel, M. Targeting DNA Repair Pathways for Cancer Treatment: What's New? *Futrure Oncol.* **2015**, *10*, 1215–1237. https://doi.org/10.2217/fon.14.60.Targeting.

(3) Cheung-Ong, K.; Giaever, G.; Nislow, C. DNA-Damaging Agents in Cancer Chemotherapy: What's New? *Chem Biol* **2013**, *20*, 648–659.

(4) Fu, D.; Calvo, J.; Samson, L. D. Balancing Repair and Tolerance of DNA Damage Caused by Alkylating Agents. *Nature* **2012**, *12*, 104–120.

(5) Yao, Y.; Dai, W. Genomic Instability and Cancer. *J Carcinog Mutagen* **2014**, *5*.

(6) Gaillard, H.; Garcia-muse, T.; Aguilera, A. Replication Stress and Cancer. *Nature* **2015**, *15*, 276–289.

(7) Brown, J. S.; Carrigan, B. O.; Jackson, S. P.; Yap, T. A. Targeting DNA Repair in Cancer: Beyond PARP Inhibitors. *Cancer Discov* **2017**, 20–38.

(8) Fan, C.; Liu, W.; Cao, H.; Wen, C.; Chen, L.; Jiang, G. O6-Methylguanine DNA Methyltransferase as a Promising Target for the Treatment of Temozolomide-Resistant Gliomas. *Cell Death Dis* **2013**, *4*, 1–8.

(9) Paranjpe, A.; Zhang, R.; Ali-osman, F.; Bobustuc, G. C.; Srivenugopal, K. S. Disulfiram Is a Direct and Potent Inhibitor of Human O6-Methylguanine-DNA Methyltransferase (MGMT) in Brain Tumor Cells and Mouse Brain and Markedly Increases the Alkylating DNA Damage. *Carcinogenesis* **2014**, *35* (3), 692–702.

(10) Matsumoto, Y.; Kim, K. Excision of Deoxyribose Phosphate Residues by DNA Polymerase β During DNA Repair. *Science* **1995**, *269*, 699–702.

(11) Canitrot, Y.; Cazaux, C.; Fréchet, M.; Bouayadi, K.; Lesca, C.; Salles, B.; Hoffmann, J. S. Overexpression of DNA Polymerase Beta in Cell Results in a Mutator Phenotype and a Decreased Sensitivity to Anticancer Drugs. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 12586–90. https://doi.org/10.1073/pnas.95.21.12586.

(12) Nickoloff, J. A.; Jones, D.; Lee, S.-H.; Williamson, E. A.; Hromas, R. Drugging the Cancers Addicted to DNA Repair. *JNCI J. Natl. Cancer Inst.* **2017**, *109* (11). https://doi.org/10.1093/jnci/djx059.

(13) Donigan, K. A.; Sun, K. W.; Nemec, A. A.; Murphy, D. L.; Cong, X.; Northrup, V.; Zelterman, D.; Sweasy, J. B. Human POLB Gene Is Mutated in High Percentage of Colorectal Tumors. *J. Biol. Chem.* **2012**, *287*, 23830–23839. https://doi.org/10.1074/jbc.M111.324947.

(14) Sobol, R. W.; Horton, J. K.; Kühn, R.; Gu, H.; Singhal, R. K.; Prasad, R.; Rajewsky, K.; Wilson, S. H. Requirement of Mammalian DNA Polymerase-Beta in Base-Excision Repair. *Nature* **1996**, *379* (6561), 183–186.

(15) Horton, J. K.; Stefanick, D. F.; Prasad, R.; Gassman, N. R.; Kedar, P. S.; Wilson, S. H. Base Excision Repair Defects Invoke Hypersensitivity to PARP Inhibition. *Mol. Cancer Res.* **2014**, *12* (August), 1128–1139. https://doi.org/10.1158/1541-7786.MCR-13-0502.

(16) Horton, J. K.; Joyce-gray, D. F.; Pachkowski, B. F.; Swenberg, J. A.; Wilson, S. H. Hypersensitivity of DNA Polymerase β Null Mouse Fibroblasts Reflects Accumulation of Cytotoxic Repair Intermediates from Site-Specific Alkyl DNA Lesions. *DNA Repair Amst* **2003**, *2*, 27–48.

(17) Li, S.; Gao, Z.; Feng, Z.; Jones, S. H.; Hecht, S. M. Plant Sterols as Selective DNA Polymerase β Lyase Inhibitors and Potentiators of Bleomycin Cytotoxicity. *Bioorg Med Chem* **2004**, *12*, 4253–4258.

(18) Deng, J.; Starck, S. R.; Sun, D.; Sabat, M.; Hecht, S. M. A New 7, 8-Euphadien-Type Triterpenoid from Brackenridgea Nitida and Bleasdalea Bleasdalei That Inhibits DNA Polymerase β. *J Nat Prod* **2000**, *63*, 1356–1360.

(19) Polosina, Y. Y.; Rosenquist, T. A.; Grollman, A. P.; Miller, H. "Knock down" of DNA Polymerase b by RNA Interference: Recapitulation of Null Phenotype. *DNA Repair* **2004**, *3*, 1469–1474. https://doi.org/10.1016/j.dnarep.2004.05.011.

(20) Barakat, K. H.; Gajewski, M. M.; Tuszynski, J. A. DNA Polymerase Beta (Pol b) Inhibitors: A Comprehensive Overview. *Drug Discov. Today* **2012**, *17*, 913–920. https://doi.org/10.1016/j.drudis.2012.04.008.

(21) Goellner, E. M.; Svilar, D.; Almeida, K. H.; Sobol, R. W. Targeting DNA Polymerase β for Therapeutic Intervention. *Curr. Mol. Pharmacol.* **2011**, *5* (1), 68–87. https://doi.org/10.2174/1874467211205010068.

(22) Mizushina, Y.; Yoshida, S.; Matsukage, a; Sakaguchi, K. The Inhibitory Action of Fatty Acids on DNA Polymerase Beta. *Biochim. Biophys. Acta* **1997**, *1336* (3), 509–21. https://doi.org/10.1016/S0304-4165(97)00067-6.

(23) Mizushina, Y.; Kamisuki, S.; Kasai, N.; Shimazaki, N.; Takemura, M.; Asahara, H.; Linn, S.; Yoshida, S.; Matsukage, A.; Koiwai, O.; Sugawara, F.; Yoshida, H.; Sakaguchi, K. A Plant Phytotoxin, Solanapyrone A, Is an Inhibitor of DNA Polymerase β and λ. *J. Biol. Chem.* **2002**, *277* (1), 630–638. https://doi.org/10.1074/jbc.M105144200.

(24) Hu, H. Y.; Horton, J. K.; Gryk, M. R.; Prasad, R.; Naron, J. M.; Sun, D. A.; Hecht, S. M.; Wilson, S. H.; Mullen, G. P. Identification of Small Molecule Synthetic Inhibitors of DNA Polymerase b by NMR Chemical Shift Mapping. *J. Biol. Chem.* **2004**, *279* (38), 39736–39744. https://doi.org/10.1074/jbc.M402842200.

(25) Gowda, A. S. P.; Suo, Z.; Spratt, T. E. Honokiol Inhibits DNA Polymerases β and λ and Increases Bleomycin Sensitivity of Human Cancer Cells. *Chem. Res. Toxicol.* **2017**, *30*, 715–725. https://doi.org/10.1021/acs.chemrestox.6b00451.

(26) Arian, D.; Hedayati, M.; Zhou, H.; Bilis, Z.; Chen, K.; Deweese, T. L.; Greenberg, M. M. Irreversible Inhibition of DNA Polymerase Beta by Small-Molecule Mimics of a DNA Lesion. *J. Am. Chem. Soc.* **2014**, *136*, 3176–3183. https://doi.org/10.1021/ja411733s.

(27) Paul, R.; Banerjee, S.; Greenberg, M. M. Synergistic Effects of an Irreversible DNA Polymerase Inhibitor and DNA Damaging Agents on HeLa Cells. *ACS Chem. Biol.* **2017**, *DOI: 10.10*.

(28) Mizushina, Y.; Murakami, C.; Yogi, K.; Ueda, K.; Ishidoh, T.; Takemura, M.; Perpelescu, M.; Suzuki, M.; Oshige, M.; Yamaguchi, T.; Saneyoshi, M.; Yoshida, H.; Sakaguchi, K. Kohamaic Acid A, a Novel Sesterterpenic Acid, Inhibits Activities of DNA Polymerases from Deuterostomes. *Biochim. Biophys. Acta - Proteins Proteomics* **2003**, *1648* (1–2), 55–61. https://doi.org/10.1016/S1570-9639(03)00108-0.

(29) Guan, L.; Greenberg, M. M. Irreversible Inhibition of DNA Polymerase Beta by an Oxidized Abasic Lesion. *J. Am. Chem. Soc.* **2010**, *132*, 5004–5005. https://doi.org/10.1021/ja101372c.

(30) Guan, L.; Bebenek, K.; Kunkel, T. A.; Greenberg, M. M. Inhibition of Short Patch and Long Patch Base Excision Repair by an Oxidized Abasic Site. *Biochemistry* **2010**, *49*, 9904–9910. https://doi.org/10.1021/bi101533a.

(31) Dimitri, A.; Burns, J. A.; Broyde, S.; Scicchitano, D. A. Transcription Elongation Past O6-Methylguanine by Human RNA Polymerase II and Bacteriophage T7 RNA Polymerase. *Nucleic Acids Res.* **2008**, *36* (20), 6459–6471.

(32) Kelland, L. The Resurgence of Platinum-Based Cancer Chemotherapy. *Nat Rev* **2007**, *7*, 573–585.

(33) Cronstein, B.; Bertino, J. R. *Methotrexate*; 2000.

(34) Longley, D. B.; Harkin, D. P.; Johnston, P.; G. 5-Fluorouracil: Mechanisms of Action and Clinical Strategies. *Nature* **2003**, *3*, 330–338.

(35) Sausville, E. A.; Peisach, J.; Horwitz, S. B. Effect of Chelating Agents and Metal Ions on the Degradation of DNA by Bleomycin. *Biochemistry* **1978**, *17* (14), 2740–2746.

(36) Burger, R. M.; Peisach, J. Activated Bleomycin. *J Biol Chem* **1981**, *256* (22), 11636–11644.

(37) Demain, A. L.; Vaishnav, P. Natural Products for Cancer Chemotherapy. *Microb Biotechnol* **2011**, *4* (6), 687–699.

(38) Rabow, L. E.; Stubbe, J.; Kozarich, J. W. Identification and Quantitation of the Lesion Accompanying Base Release in Bleomycin-Mediated DNA Degradation. *J Am Chem Soc* **1990**, *112*, 3196–3203.

(39) Pommier, Y. Topoisomerase I Inhibitors: Camptothecins and Beyong. *Nature* **2006**, *6*, 789–802.

(40) van Vuuren, R.; H Visagie, M.; Theron, A.; Joubert, A. Antimitotic Drugs in the Treatment of Cancer. *Cancer Chemother Pharmacol* **2015**, *76*, 1101–1112.

(41) Hoeijmakers, J. H. J. DNA Damage, Aging, and Cancer. *N Engl J Med* **2009**, *361* (15), 1475–1485.

(42) Greenberg, M. M. Elucidating DNA Damage and Repair Processes by Independently Generating Reactive and Metastable Intermediates. *Org. Biomol. Chem.* **2007**, *5*, 18–30. https://doi.org/10.1039/b612729k.

(43) Greenberg, M. M. Looking beneath the Surface to Determine What Makes DNA Damage Deleterious. *Curr Opin Chem Biol* **2014**, *21*, 48–55.

(44) Lindahl, T.; Nyberg, B. Rate of Depurination of Native Deoxyribonucleic Acid. *Biochemistry* **1972**, *11* (19), 3610–3618.

(45) Kubo, K.; Ide, H.; Wallace, S. S.; Kow, Y. W. A Novel Sensitive and Specific Assay for Abasic Sites, the Most Commonly Produced DNA Lesion. *Biochemistry* **1992**, *31*, 3703–3708.

(46) Dutta, S.; Chowdhury, G.; Gates, K. S. Interstrand Cross-Links Generated by Abasic Sites in Duplex DNA. *J. Am. Chem. Soc.* **2007**, *129*, 1852–1853. https://doi.org/10.1021/ja067294u.

(47) Price, N. E.; Johnson, K. M.; Wang, J.; Fekry, M. I.; Wang, Y.; Gates, K. S. Interstrand DNA-DNA Cross-Link Formation between Adenine Residues and Abasic Sites in Duplex Dna. *J. Am. Chem. Soc.* **2014**, *136*, 3483–3490. https://doi.org/10.1021/ja410969x.

(48) Beger, R. D.; Bolton, P. H. Structures of Apurinic and Apyrimidinic Sites in Duplex DNAs. *J Biol Chem* **1998**, *273* (25), 15565–15573.

(49) Luke, A. M.; Chastain, P. D.; Pachkowski, B. F.; Afonin, V.; Takeda, S.; Kaufman, D. G.; Swenberg, J. A.; Nakamura, J. Accumulation of True Single Strand Breaks and AP Sites in Base Excision Repair Deficient Cells. *Mutat. Res.* **2010**, *694*, 65–71.

(50) Chen, H.; Cui, Z.; Hejazi, L.; Yao, L.; Walmsley, S. J.; Rizzo, C. J.; Turesky, R. J. Kinetics of DNA Adducts and Abasic Site Formation in Tissues of Mice Treated with a Nitrogen Mustard. *Chem. Res. Toxicol.* **2020**.

(51) Murphy, M. P. How Mitochondria Produce Reactive Oxygen Species. *Biochem J* **2009**, *13*, 1–13.

(52) Hart, P. J.; Balbirnie, M. M.; Ogihara, N. L.; Nersissian, A. M.; Weiss, M. S.; Valentine, J. S.; Eisenberg, D. A Structure-Based Mechanism for Copper - Zinc Superoxide Dismutase. *Biochemistry* **1999**, *38*, 2167–2178.

(53) Winterbourn, C. C. Toxicity of Iron and Hydrogen Peroxide: The Fenton Reaction. *Toxicol Lett* **1995**, *83*, 969–974.

(54) Laverne, J. A. OH Radicals and Oxidizing Products in the Gamma Radiolysis of Water. *Radiat Res* **2000**, *153* (2), 196–200.

(55) von. Sonntag, C. *Free-Radical-Induced DNA Damage and Its Repair*; 2006.

(56) Balasubramanian, B.; Pogozelski, W.; Tullius, T. DNA Strand Breaking by the Hydroxyl Radical Is Governed by the Accessible Surface Areas of the Hydrogen Atoms of the DNA Backbone. *Proc. Natl. Acad. Sci.* **1998**, *95*, 9738–9743.

(57) Avendano, C.; Menendez, J. C. *Medicinal Chemistry of Anticancer Drugs*; 2008.

(58) Chen, B.; Bohnert, T.; Zhou, X.; Dedon, P. C. 5′-(2-Phosphoryl-1,4-Dioxobutane) as a Product of 5′-Oxidation of Deoxyribose in DNA: Elimination as Trans-1,4-Dioxo-2-Butene and Approaches to Analysis. *Chem. Res. Toxicol.* **2004**, *17*, 1406–1413.

(59) Demott, M. S.; Beyret, E.; Wong, D.; Bales, B. C.; Hwang, J. T.; Greenberg, M. M.; Demple, B. Covalent Trapping of Human DNA Polymerase b by the Oxidative DNA Lesion 2-Deoxyribonolactone. *J. Biol. Chem.* **2002**, *277*, 7637–7640. https://doi.org/10.1074/jbc.C100577200.

(60) Greenberg, M. M. Abasic and Oxidized Abasic Site Reactivity in DNA: Enzyme Inhibition, Cross-Linking, and Nucleosome Catalyzed Reactions. *Acc. Chem. Res.* **2014**, *47*, 646–655.

(61) Guan, L.; Greenberg, M. M. DNA Interstrand Cross-Link Formation by the 1,4-Dioxobutane Abasic Lesion. *J Am Chem Soc* **2009**, *131*, 15225–15231.

(62) Byrns, M. C.; Predecki, D. P.; Peterson, L. A. Characterization of Nucleoside Adducts of Cis-2-Butene-1,4-Dial, a Reactive Metabolite of Furan. *Chem. Res. Toxicol.* **2002**, *15*, 373–379. https://doi.org/10.1021/tx0101402.

(63) Chen, B.; Vu, C. C.; Byrns, M. C.; Dedon, P. C.; Peterson, L. A. Formation of 1,4-Dioxo-2-Butene-Derived Adducts of 2'-Deoxyadenosine and 2'Deoxycytidine in Oxidized DNA. *Chem. Res. Toxicol.* **2006**, *19*, 982–985.

(64) Guan, L.; Greenberg, M. M. An Oxidized Abasic Lesion as an Intramolecular Source of DNA Adducts. *Aust J Chem* **2011**, *64*, 438–442.

(65) Weng, L.; Greenberg, M. M. Rapid Histone-Catalyzed DNA Lesion Excision and Accompanying Protein Modification in Nucleosomes and Nucleosome Core Particles. *J Am Chem Soc* **2015**, *137*, 11022–11031.

(66) Zhou, C.; Sczepanski, J. T.; Greenberg, M. M. Histone Modification via Rapid Cleavage of C4′-Oxidized Abasic Sites in Nucleosome Core Particles. *J Am Chem Soc* **2013**, *135*, 5274–5277.

(67) Regulus, P.; Duroux, B.; Bayle, P.-A.; Favier, A.; Cadet, J.; Ravanat, J.-L. Oxidation of the Sugar Moiety of DNA by Ionizing Radiation or Bleomycin Could Induce the Formation of a Cluster DNA Lesion. *Proc. Natl. Acad. Sci.* **2007**, *104*, 14032–14037.

(68) Sczepanski, J. T.; Jacobs, A. C.; Van Houten, B.; Greenberg, M. M. Double-Strand Break Formation during Nucleotide Excision Repair of a DNA Interstrand Cross-Link. *Biochemistry* **2009**, *48*, 7565–7567. https://doi.org/10.1021/bi901006b.

(69) Mentegari, E.; Kissova, M.; Bavagnoli, L.; Maga, G.; Crespan, E. DNA Polymerases λ and B : The Double-Edged Swords of DNA Repair. *Genes* **2016**, *7*, 1–17. https://doi.org/10.3390/genes7090057.

(70) Barakat, K. DNA Repair Inhibitors: The Next Major Step to Improve Cancer Therapy. *Curr. Top. Med. Chem.* **2012**, *12*, 1376–1390. https://doi.org/10.2174/156802612801319070.

(71) Lindahl, T. Past, Present, and Future Aspects of Base Excision Reapir. *Prog Nucleic Acid Res Mol Biol* **2001**, *68*, 17–30.

(72) Dianov, G. L.; Hu, U. Mammalian Base Excision Repair: The Forgotten Archangel. *Nucleic Acids Res.* **2013**, *41* (6), 3483–3490.

(73) Srivastava, D. K.; Berg, B. J. Vande; Prasad, R.; Molina, J. T.; Beard, W. A.; Tomkinson, A. E.; Wilson, S. H. Mammalian Abasic Site Base Excision Repair. *J Biol Chem* **1998**, *273* (33), 21203–21209.

(74) Krokan, H. E.; Bjoras, M. Base Excision Repair. *Cold Spring Harb Perspect Biol* **2013**, *5*, 1–22.

(75) McCullough, A. K.; Dodson, M. L.; Lloyd, R. S. Initiation of Base Excision Repair: Glycosylase. *Ann Rev Biochem* **1999**, *68*, 255–285.

(76) Asagoshi, K.; Yamada, T.; Terato, H.; Ohyama, Y.; Monden, Y.; Arai, T.; Nishimura, S.; Aburatani, H.; Lindahl, T.; Ide, H. Distinct Repair Activities of Human 7, 8-Dihydro-8-Oxoguanine DNA Glycosylase and Formamidopyrimidine DNA Glycosylase for Formamidopyrimidine and 7, 8-Dihydro-8-Oxoguanine*. *J Biol Chem* **2000**, *275* (7), 4956–4964.

(77) Takao, M.; Kanno, S.; Shiromoto, T.; Hasegawa, R.; Ide, H.; Ikeda, S.; Sarker, A. H.; Seki, S.; Xing, J. Z.; Le, X. C.; Weinfeld, M.; Kobayashi, K.; Miyazaki, J.; Muijtjens, M.; Hoeijmakers, J. H. J.; Horst, G.; Yasui, A. Novel Nuclear and Mitochondrial Glycosylases Revealed by Disruption of the Mouse Nth1 Gene Encoding an Endonuclease III Homolog for Repair of Thymine Glycols. *EMBO J* **2002**, *21* (12), 3486–3493.

(78) Sobol, R. W.; Wilson, S. H. Mammalian DNA, β-Polymerase in Base Excision Repair of Alkylation Damage. *Prog Nucleic Acid Res Mol Biol* **2001**, *68*, 57–74.

(79) Manoharan, Muthiah.; Mazumder, Abhijit.; Ransom, S. C.; Gerlt, J. a; Bolton, P. H. Mechanism of UV Endonuclease V Cleavage of Abasic Sites in DNA Determined by Carbon-13 Labeling. *J. Am. Chem. Soc.* **1988**, *110*, 2690–2691. https://doi.org/10.1021/ja00216a074.

(80) Prasad, R.; Beard, W. A.; Chyan, J. Y.; Maciejewski, M. W.; Mullen, G. P.; Wilson, S. H. Functional Analysis of the Amino-Terminal 8-KDa Domain of DNA Polymerase B as Revealed by Site-Directed Mutagenesis. *J. Biol. Chem.* **1998**, *273*, 11121–11126.

(81) Prasad, R.; Beard, W. A.; Wilson, S. H. Studies of Gapped DNA Substrate Binding by Mammalian DNA Polymerase β Dependence on 5'-Phosphate Group. *J. Biol. Chem.* **1994**, *269*, 18096–18101.

(82) Wilson, S. H.; Beard, W. A.; Shock, D. D.; Batra, V. K. Base Excision Repair and Design of Small Molecule Inhibitors of Human DNA Polymerase b. *Cell Mol Life Sci* **2010**, *67*, 3633–3647. https://doi.org/10.1007/s00018-010-0489-1.

(83) Krokan, H. E.; Standal, R.; Slupphaug, G. DNA Glycosylases in the Base Excision Repair of DNA. *Biochem J* **1997**, *325*, 1–16.

(84) Da, L.; Yu, J. Base-Flipping Dynamics from an Intrahelical to an Extrahelical State Exerted by Thymine DNA Glycosylase during DNA Repair Process. *Nucleic Acids Res.* **2018**, *46* (11), 5410–5425.

(85) Maher, R. L.; Wallace, S. S.; Pederson, D. S. The Lyase Activity of Bifunctional DNA Glycosylases and the 3'-Diesterase Activity of APE1 Contribute to the Repair of Oxidized Bases in Nucleosomes. *Nucleic Acids Res.* **2019**, *47* (6), 2922–2931.

(86) Hill, J. W.; Hazra, T. K.; Izumi, T.; Mitra, S. Stimulation of Human 8-Oxoguanine-DNA Glycosylase by AP-Endonuclease: Potential Coordination of the Initial Steps in Base Excision Repair. *Nucleic Acids Res.* **2001**, *29* (2), 430–438.

(87) Albelazi, M.; Martin, P.; Mohammed, S.; Mutti, L.; Parsons, J. L.; Elder, R. H. The Biochemical Role of the Human NEIL1 and NEIL3 DNA Glycosylases on Model DNA Replication Forks. *Genes Basel* **2019**, *10* (315), 1–11.

(88) Takao, M.; Oohata, Y.; Kitadokoro, K.; Kobayashi, K.; Iwai, S.; Yasui, A.; Yonei, S.; Zhang, Q. Human Nei-like Protein NEIL3 Has AP Lyase Activity Specific for Single-Stranded DNA and Confers Oxidative Stress Resistance in Escherichia Coli Mutant. *Genes Cells* **2009**, *14*, 261–270.

(89) Maher, R. L.; Bloom, L. B. Pre-Steady-State Kinetic Characterization of the AP Endonuclease Activity of Human AP Endonuclease 1. *J Biol Chem* **2007**, *282* (42), 30577–30585.

(90) Xu, Y.; Demott, M. S.; Taeg, J.; Greenberg, M. M.; Demple, B. Action of Human Apurinic Endonuclease (Ape1) on C1'-Oxidized Deoxyribose Damage in DNA. *DNA Repair Amst* **2003**, *2*, 175–185.

(91) Jacobs, A. C.; Kreller, C. R.; Greenberg, M. M. Long Patch Base Excision Repair Compensates for DNA Polymerase b Inactivation by the C4'-Oxidized Abasic Site. *Biochemistry* **2011**, *50*, 136–143. https://doi.org/10.1021/bi1017667.

(92) Xanthoudakis, S.; Smeyne, R. J.; Wallace, J. D.; Curran, T. The Redox/DNA Repair Protein, Ref-1, Is Essential for Early Embryonic Development in Mice. *Proc. Natl. Acad. Sci.* **1996**, *93*, 8919–8923.

(93) Ludwig, D. L.; MacInnes, M. A.; Takiguchi, Y.; Purtymun, P. E.; Henrie, M.; Flannery, M.; Meneses, J.; Pedersen, R. A.; Chen, D. J. A Murine AP-Endonuclease Gene-Targeted Deficiency with Post-Implantation Embryonic Progression and Ionizing Radiation Sensitivity. *Mutat. Res. - DNA Repair* **1998**, *409*, 17–29.

(94) Lange, S. S.; Takata, K.; Wood, R. D. DNA Polymerases and Cancer. *Nat. Rev. Cancer* **2011**, *11* (2), 96–110.

(95) Bebenek, K.; Tissier, A.; Frank, E. G.; Mcdonald, J. P.; Prasad, R.; Wilson, S. H.; Woodgate, R.; Kunkel, T. A. 5'-Deoxyribose Phosphate Lyase Activity of Human DNA Polymerase I in Vitro. *Science* **2001**, *291* (5511), 2156–2159.

(96) Prasad, R.; Longley, M. J.; Sharief, F. S.; Hou, E. W.; Copeland, W. C.; Wilson, S. H. Human DNA Polymerase Q Possesses 5'-DRP Lyase Activity and Functions in Single-Nucleotide Base Excision Repair in Vitro. *Nucleic Acids Res.* **2009**, *37* (6), 1868–1877.

(97) Yamtich, J.; Sweasy, J. B. DNA Polymerase Family X: Function, Structure, and Cellular Roles. *Biochim. Biophys. Acta* **2010**, *1804* (5), 1136–1150.

(98) Garcia-Diaz, M.; Dominguez, O.; Lopez-Fernandez, L. A.; Lera, L. T. de; Saniger, M. L.; Ruiz, J. F.; Parraga, M.; Garcia-Ortiz, M. J.; Kirchhoff, T.; Mazo, J. del; Bernad, A.; Blanco, L. DNA Polymerase Lambda (Pol L), a Novel Eukaryotic DNA Polymerase with a Potential Role in Meiosis. *J Mol Biol* **2000**, *301*, 851–867.

(99) Belousova, E. A.; Lavrik, O. I. DNA Polymerases B and L and Their Roles in Cell. *DNA Repair Amst* **2015**, *29*, 112–126.

(100) Brown, J. A.; Pack, L. R.; Sherrer, S. M.; Kshetry, A. K.; Newmister, S. A.; Fowler, J. D.; Taylor, J.; Suo, Z. Identification of Critical Residues for the Tight Binding of Both Correct and Incorrect Nucleotides to Human DNA Polymerase λ. *J Mol Biol* **2010**, *403*, 505–515.

(101) Starcevic, D.; Dalal, S.; Sweasy, J. B. Is There a Link between DNA Polymerase Beta and Cancer? *Cell Cycle* **2004**, *3*, 998–1001. https://doi.org/10.4161/cc.3.8.1062.

(102) Braithwaite, E. K.; Prasad, R.; Shock, D. D.; Hou, E. W.; Beard, W. A.; Wilson, S. H. DNA Polymerase b Mediates a Back-up Base Excision Repair Activity in Extracts of Mouse Embryonic Fibroblasts. *J. Biol. Chem.* **2005**, *280*, 18469–18475. https://doi.org/10.1074/jbc.M411864200.

(103) Piersen, C. E.; Prasad, R.; Wilson, S. H.; Lloyd, R. S. Evidence for an Imino Intermediate in the DNA Polymerase β Deoxyribose Phosphate Excision Reaction. *J. Biol. Chem.* **1996**, *271*, 17811–17815. https://doi.org/10.1074/jbc.271.30.17811.

(104) Singhal, R. K.; Wilson, S. H. Short Gap-Filling Synthesis by DNA Polymerase Beta Is Processive. *J. Biol. Chem.* **1993**, *268*, 15906–11.

(105) Feng, J.; Crasto, C. J.; Matsumoto, Y.; April, R. V.; Re, V.; Recei, M.; May, V. Deoxyribose Phosphate Excision by the N-Terminal Domain of the Polymerase : *Biochemistry* **1998**, *37*, 9605–9611.

(106) Mullen, G. P.; Wilson, S. H. DNA Polymerase b in Abasic Site Repair: A Structurally Conserved Helix- Hairpin-Helix Motif in Lesion Detection by Base Excision Repair Enzymes. *Biochemistry* **1997**, *36*, 4713–4717. https://doi.org/10.1021/bi962363a.

(107) Matsumoto, Y.; Kim, K.; Katz, D. S.; Feng, J. A. Catalytic Center of DNA Polymerase b for Excision of Deoxyribose Phosphate Groups. *Biochemistry* **1998**, *37*, 6456–6464. https://doi.org/10.1021/bi9727545.

(108) Prasad, R.; Batra, V. K.; Yang, X. P.; Krahn, J. M.; Pedersen, L. C.; Beard, W. A.; Wilson, S. H. Structural Insight into the DNA Polymerase β Deoxyribose Phosphate Lyase Mechanism. *DNA Repair Amst* **2005**, *4*, 1347–1357.

(109) Fromme, J. C.; Verdine, G. L. Structure of a Trapped Endonuclease III ± DNA Covalent Intermediate. *Eur. Mol. Biol. Organ. J.* **2003**, *22*, 3461–3471.

(110) Pelletier, H.; Sawaya, M. R.; Kumar, A.; Wilson, S. H.; Kraut, J. Structures of Ternary Complexes of Rat DNA Polymerase Beta, a DNA Template-Primer, and DdCTP. *Science* **1994**, *264* (5167), 1891–903. https://doi.org/066/1.

(111) Kumar, A.; Abbotts, J.; Karawya, E. M.; Wilson, S. H. Identification and Properties of the Catalytic Domain of Mammalian DNA Polymerase β. *Biochemistry* **1990**, *29*, 7156–7159.

(112) Pelletier, H.; Sawaya, M. R.; Wolfle, W.; Wilson, S. H.; Kraut, J. A Structural Basis for Metal Ion Mutagenicity and Nucleotide Selectivity in Human DNA Polymerase β. *Biochemistry* **1996**, *35* (39), 12762–12777. https://doi.org/10.1021/bi9529566.

(113) Kunkel, T. A. DNA Replication Fidelity. *J Biol Chem* **2004**, *279* (17), 16895–16898.

(114) Pelletier, H.; Sawaya, M. R.; Wolfle, W.; Wilson, S. H.; Kraut, J. Crystal Structures of Human DNA Polymerase β Complexed with DNA: Implications for Catalytic Mechanism, Processivity, and Fidelity. *Biochemistry* **1996**, *35*, 12742–12761.

(115) Sawaya, M. R.; Pelletier, H.; Kumar, A.; Wilson, S. H.; Sawaya, M. R.; Pelletier, H.; Kumar, A.; Wilson, S. H.; Kraut, J. Crystal Structure of Rat DNA Polymerase B : Evidence for a Common Polymerase Mechanism. **1994**, *264* (5167), 1930–1935.

(116) Sawaya, M. R.; Prasad, R.; Wilson, S. H.; Kraut, J.; Pelletier, H. Crystal Structures of Human DNA Polymerase β Complexed with Gapped and Nicked DNA: Evidence for an Induced Fit Mechanism † , ‡. *Biochemistry* **1997**, *36* (37), 11205–11215. https://doi.org/10.1021/bi9703812.

(117) Yang, L.; Beard, W. A.; Wilson, S. H.; Broyde, S.; Schlick, T. Highly Organized but Pliant Active Site of DNA Polymerase β: Compensatory Mechanisms in Mutant Enzymes Revealed by Dynamics Simulations and Energy Analyses. *Biophys J* **2004**, *86*, 3392–3408.

(118) Kraynov, V. S.; Showalter, A. K.; Liu, J.; Zhong, X.; Tsai, M. DNA Polymerase β: Contributions of Template-Positioning and DNTP Triphosphate-Binding Residues to Catalysis and Fidelity. *Biochemistry* **2000**, *39*, 16008–16015.

(119) Steensel, B. Van; Smogorzewska, A.; de Lange, T. TRF2 Protects Human Telomeres from End-to-End Fusions. *Cell* **1998**, *92*, 401–413.

(120) Menck, C. F. M.; Munford, V. DNA Repair Diseases: What Do They Tell Us about Cancer and Aging? *Genet Mol Biol* **2014**, *37*, 220–233.

(121) Chen, K. H.; Yakes, F. M.; Srivastava, D. K.; Singhal, R. K.; Sobol, R. W.; Horton, J. K.; Van Houten, B.; Wilson, S. H. Up-Regulation of Base Excision Repair Correlates with Enhanced Protection against a DNA Damaging Agent in Mouse Cell Lines. *Nucleic Acids Res.* **1998**, *26* (8), 2001–2007. https://doi.org/10.1093/nar/26.8.2001.

(122) Hong, Z.; Peng, X.; Min, L.; Jimin, Z.; Ziming, D.; Guoqiang, Z. DNA Polymerase Beta Overexpression Correlates with Poor Prognosis in Esophageal Cancer Patients. *Chin. Sci Bull* **2013**, *58* (26), 3274–3279.

(123) Abdel-fatah, T. M. A.; Russell, R.; Agarwal, D.; Moseley, P.; Abayomi, M.; Perry, C.; Albarakati, N.; Ball, G.; Chan, S.; Caldas, C.; Ellis, I. O.; Madhusudan, S. DNA Polymerase β Deficiency Is Linked to Aggressive Breast Cancer: A Comprehensive Analysis of Gene Copy Number, MRNA and Protein Expression in Multiple Cohorts. *Mol Oncol* **2014**, *8*, 520–532.

(124) Husain, I.; Morton, B. S.; Beard, W. a; Singhal, R. K.; Prasad, R.; Wilson, S. H.; Besterman, J. M. Specific Inhibition of DNA Polymerase Beta by Its 14 KDa Domain: Role of Single- and Double-Stranded DNA Binding and 5'-Phosphate Recognition. *Nucleic Acids Res.* **1995**, *23* (9), 1597–1603. https://doi.org/4a0776 [pii].

(125) Liu, S.; Lai, Y.; Zhao, W.; Wu, M.; Zhang, Z. Links between DNA Polymerase Beta Expression and Sensitivity to Bleomycin. *Toxicology* **2011**, *281*, 63–69.

(126) Horton, J. K.; Srivastava, D. K.; Zmudzka, B. Z.; Wilson, S. H. Strategic Down-Regulation of DNA Polymerase β by Antisense RNA Sensitizes Mammalian Cells to Specific DNA Damaging Agents. *Nucleic Acids Res.* **1995**, *23* (19), 3810–3815.

(127) Sugo, N.; Aratani, Y.; Nagashima, Y.; Kubota, Y.; Koyama, H. Neonatal Lethality with Abnormal Neurogenesis in Mice Deficient in DNA Polymerase Beta. *EMBO J.* **2000**, *19*, 1397–404. https://doi.org/10.1093/emboj/19.6.1397.

(128) Gu, H.; Marth, J. D.; Orban, P. C.; Mossmann, H.; Rajewsky, K. Deletion of a DNA Polymerase IP Gene Segment in T Cells Using Cell Type-Specific Gene Targeting Type Deletion. *Science* **1992**, *265*, 103–106.

(129) Nemec, A. A.; Donigan, K. A.; Murphy, D. L.; Jaegers, J.; Sweasy, J. B. Colon Cancer-Associated DNA Polymerase β Variant Induces Genomic Instability and Cellular Transformation. *J. Biol. Chem.* **2012**, *287*, 23840–23849. https://doi.org/10.1074/jbc.M112.362111.

(130) Tan, X.; Wang, H.; Luo, G.; Ren, S.; Li, W.; Cui, J.; Gill, H. S.; Fu, S. W.; Lu, Y. Clinical Significance of a Point Mutation in DNA Polymerase Beta (POLB) Gene in Gastric Cancer. *Int. J. Biol. Sci.* **2015**, *11* (2), 144–155. https://doi.org/10.7150/ijbs.10692.

(131) Khanra, K.; Panda, K.; Bhattacharya, C.; Mitra, A. K.; Sarkar, R. Association between Newly Identi Fi Ed Variant Form of DNA Polymerase Beta Δ208 − 304 and Ovarian Cancer. **2012**, *11*, 155–160. https://doi.org/10.3233/CBM-2012-0275.

(132) Sung, J.; Demple, B. Roles of Base Excision Repair Subpathways in Correcting Oxidized Abasic Sites in DNA. *FEBS J* **2006**, *273*, 1620–1629.

(133) Beaver, J. M.; Lai, Y.; Rolle, S. J.; Weng, L.; Greenberg, M. M.; Liu, Y. An Oxidized Abasic Lesion Inhibits Base Excision Repair Leading to DNA Strand Breaks in a Trinucleotide Repeat Tract. *PLoS One* **2018**, *1*, 1–21.

(134) Stevens, A. J.; Guan, L.; Bebenek, K.; Kunkel, T. A.; Greenberg, M. M. DNA Polymerase λ Inactivation by Oxidized Abasic Sites. *Biochemistry* **2013**, *52*, 975–983. https://doi.org/10.1021/bi301592x.

(135) Laverty, D. J.; Mortimer, I. P.; Greenberg, M. M. Mechanistic Insight through Irreversible Inhibition: DNA Polymerase $þeta$ Uses a Common Active Site for Polymerase and Lyase Activities. *J Am Chem Soc* **2018**, *140*, 9034–9037.

(136) Antonarakis, E. S.; Gomella, L. G.; Petrylak, D. P. When and How to Use PARP Inhibitors in Prostate Cancer: A Systematic Review of the Literature with an Update on On-Going Trials. *Eur. Urol. Oncol.* **2020**, *3* (5), 594–611. https://doi.org/10.1016/j.euo.2020.07.005.

(137) To, C.; Kim, E.-H.; Royce, D. B.; Williams, C. R.; Collins, R. M.; Risingsong, R.; Sporn, M. B.; Liby, K. T. The PARP Inhibitors, Veliparib and Olaparib, Are Effective Chemopreventive Agents for Delaying Mammary Tumor Development in BRCA1-Deficient Mice. *Cancer Prev. Res. (Phila. Pa.)* 11.

(138) Herceg, Z.; Wang, Z. Functions of Poly (ADP-Ribose) Polymerase (PARP) in DNA Repair, Genomic Integrity and Cell Death. *Mutat. Res.* **2001**, *477*, 97–110.

(139) Murai, J.; Huang, S. Y. N.; Das, B. B.; Renaud, A.; Zhang, Y.; Doroshow, J. H.; Ji, J.; Takeda, S.; Pommier, Y. Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors. *Cancer Res.* **2012**, *72*, 5588–5599. https://doi.org/10.1158/0008-5472.CAN-12-2753.

(140) Dziadkowiec, K. N.; G\c asiorowska, E.; Nowak-markwitz, E.; Jankowska, A. PARP Inhibitors: Review of Mechanisms of Action and BRCA1/2 Mutation Targeting. *Menopause Rev* **2016**, *15* (4), 215–219.

(141) Gavande, N. S.; VanderVere-Carozza, P. S.; Hinshaw, H. D.; Jalal, S. I.; Sears, C. R.; Pawelczak, K. S.; Turchi, J. J. DNA Repair Targeted Therapy: The Past or Future of

Cancer Treatment? *Pharmacol. Ther.* **2016**, *160*, 65–83. https://doi.org/10.1016/j.pharmthera.2016.02.003.

(142) Mizushina, Y.; Tanaka, N.; Yagi, H.; Kurosawa, T.; Onoue, M.; Seto, H.; Horie, T.; Aoyagi, N.; Yamaoka, M.; Matsukage, A.; Yoshida, S.; Sakaguchi, K. Fatty Acids Selectively Inhibit Eukaryotic DNA Polymerase Activities in Vitro. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1996**, *1308* (3), 256–262. https://doi.org/10.1016/0167-4781(96)00121-2.

(143) Mizushina, Y.; Sagisaka, M.; Sakaib, H.; Abeb, M.; Sakaguchi, K. Mode Analysis of Binding of Fatty Acids to Mammalian DNA Polymerases. *Biochim. Biophys. Acta* **2000**, *1486* (2–3), 211–8.

(144) Sun, D. A.; Starck, S. R.; Locke, E. P.; Hecht, S. M. DNA Polymerase β Inhibitors from Sandoricum Koetjape. *J. Nat. Prod.* **1999**, *62* (8), 1110–1113. https://doi.org/10.1021/np990104r.

(145) Hazan, C.; Boudsocq, F.; Gervais, V.; Saurel, O.; Ciais, M.; Cazaux, C.; Czaplicki, J.; Milon, A. Structural Insights on the Pamoic Acid and the 8 KDa Domain of DNA Polymerase Beta Complex: Towards the Design of Higher-Affinity Inhibitors. *BMC Struct. Biol.* **2008**, *8*, 22. https://doi.org/10.1186/1472-6807-8-22.

(146) Barakat, K.; Tuszynski, J. Relaxed Complex Scheme Suggests Novel Inhibitors for the Lyase Activity of DNA Polymerase Beta. *J. Mol. Graph. Model.* **2011**, *29* (5), 702–716. https://doi.org/10.1016/j.jmgm.2010.12.003.

(147) Ogawa, A.; Murate, T.; Suzuki, M.; Nimura, Y.; Yoshida, S. Lithocholic Acid, a Putative Tumor Promoter, Inhibits Mammalian DNA Polymerase Beta. *Jpn. J. Cancer Res. Gann* **1998**, *89*, 1154–1159.

(148) Nguyen, T.; Ung, T.; Jung, Y. Role of Bile Acids in Colon Carcinogenesis. *World J Clin Cases* **2018**, *6* (13), 577–589.

(149) Mooring, S. R.; Jin, H.; Devi, N. S.; Jabbar, A. A.; Kaluz, S.; Liu, Y.; Van Meir, E. G.; Wang, B. Design and Synthesis of Novel Small-Molecule Inhibitors of the Hypoxia Inducible Factor Pathway. *J. Med. Chem.* **2011**, *54* (24), 8471–8489. https://doi.org/10.1021/jm201018g.

(150) Lenci, E.; Trabocchi, A. Synthetic Approaches toward Small Molecule Libraries. In *Small Molecule Drug Discovery*; Elsevier, 2020; pp 1–34. https://doi.org/10.1016/B978-0-12-818349-6.00001-7.

(151) Riel-Mehan, M.; Shokat, K. A Crosslinker Based on a Tethered Electrophile for Mapping Kinase Substrate Networks. *Cell Chem. Biol.* **2014**, *21* (5), 585–590.

(152) Mootoo, D.; Konradsson, P.; Udodong, U.; Fraser-Reid, B. Armed and Disarmed N-Pentenyl Glycosides in Saccharides Couplings Leading to Oligosaccharides. *J Am Chem Soc* **1988**, *110* (16), 5582–5584.

(153) Mootoo, D.; Date, V.; Fraser-Reid, B. N-Pentenyl Glycosides Permit the Chemospecific Liberation of Anomeric Center. *J Am Chem Soc* **1988**, *110* (8), 2662–2663.

(154) Arasappan, A.; Fraser-Reid, B. N-Pentenyl Furanosides: Synthesis and Glycosidation Reactions of Some Galacto Derivatives. *Tetrahedron Lett.* **1995**, *36* (44), 7967–7970.

(155) Lu, J.; Fraser-Reid, B. Preparation, Properties, and Applications of n-Pentenyl Arabinofuranosyl Donors. *Org. Lett.* **2004**, *6* (18), 3051–3054.

(156) Hansen, A.; Thalhammer, A.; El-saheer, A.; Brown, T.; Schofield, C. Improved Synthesis of 5-Hydroxymethyl-2'-Deoxycytidine Phosphoramidite Using a 2'- Deoxyuridine to 2'-

Deoxycytidine Conversion without Temporary Protecting Groups. *Bioorg Med Chem Lett* **2011**, *21* (4), 1181–1184.

(157) Adam, J.; Gosselain, P.; Goldfinger, P. Laws of Addition and Substitution in Atomic Reactions of Halogens. *Nature* **1953**, *171*, 704–705.

(158) Dorjsuren, D.; Wilson, D. M.; Beard, W. A.; McDonald, J. P.; Austin, C. P.; Woodgate, R.; Wilson, S. H.; Simeonov, A. A Real-Time Fluorescence Method for Enzymatic Characterization of Specialized Human DNA Polymerases. *Nucleic Acids Res.* **2009**, *37* (19). https://doi.org/10.1093/nar/gkp641.

(159) Kitz, R.; Wilson, I. Esters of Methanesulfonic Acid as Irreversible Inhibitors of Acetylcholinesterase. *J. Biol. Chem.* **1962**, *237* (10), 3245–3249.

(160) Wilson, D. L.; Kool, E. T. Ultrafast Oxime Formation Enables Efficient Fluorescence Light-up Measurement of DNA Base Excision. *J Am Chem Soc* **2019**, *141*, 19379–19388.

(161) He, P.; Yang, W. Template and Primer Requirements for DNA Pol θ-Mediated End Joining. *Proc. Natl. Acad. Sci.* **2018**, *115* (30), 7747–7752. https://doi.org/10.1073/pnas.1807329115.

(162) Sebesta, M. Reconstitution of DNA Repair Synthesis in Vitro and the Role of Polymerase and Helicase Activities. *DNA Repair* **2011**, 10.

(163) Asagoshi, K.; Lehmann, W.; Braithwaite, E. K.; Santana-Santos, L.; Prasad, R.; Freedman, J. H.; Van Houten, B.; Wilson, S. H. Single-Nucleotide Base Excision Repair DNA Polymerase Activity in C. Elegans in the Absence of DNA Polymerase. *Nucleic Acids Res.* **2012**, *40* (2), 670–681. https://doi.org/10.1093/nar/gkr727.

(164) Shan, Y.; Liu, F.; Ma, Y.; Fan, T.; Wang, M.; Dong, Y. DESIGN , SYNTHESIS AND BIOLOGICAL EVALUATION OF 2-PHENOXY N-PHENYLACETAMIDES AS NOVEL ANTICANCER AGENTS. *Pharm Chem J* **2016**, *49* (12), 798–803.

(165) Daskalova, S. M.; Eisenhauer, B. M.; Gao, M.; Feng, X.; Ji, X.; Cheng, Q.; Fahmi, N.; Khdour, O. M.; Chen, S.; Hecht, S. M. An Assay for DNA Polymerase β Lyase Inhibitors That Engage the Catalytic Nucleophile for Binding. *Bioorg. Med. Chem.* **2020**, *28* (17), 115642. https://doi.org/10.1016/j.bmc.2020.115642.

(166) Chen, H.; Coseno, M.; Ficarro, S. B.; Mansueto, M. S.; Komazin-Meredith, G.; Boissel, S.; Filman, D. J.; Marto, J. A.; Hogle, J. M.; Coen, D. M. A Small Covalent Allosteric Inhibitor of Human Cytomegalovirus DNA Polymerase Subunit Interactions. *ACS Infect. Dis.* **2017**, *3* (2), 112–118. https://doi.org/10.1021/acsinfecdis.6b00079.

(167) Shunmugam, L.; Soliman, M. E. S. Targeting HCV Polymerase: A Structural and Dynamic Perspective into the Mechanism of Selective Covalent Inhibition. *RSC Adv.* **2018**, *8* (73), 42210–42222. https://doi.org/10.1039/C8RA07346E.

(168) Huang, H.; Howard, C. A.; Zari, S.; Cho, H. J.; Shukla, S.; Li, H.; Ndoj, J.; González-Alonso, P.; Nikolaidis, C.; Abbott, J.; Rogawski, D. S.; Potopnyk, M. A.; Kempinska, K.; Miao, H.; Purohit, T.; Henderson, A.; Mapp, A.; Sulis, M. L.; Ferrando, A.; Grembecka, J.; Cierpicki, T. Covalent Inhibition of NSD1 Histone Methyltransferase. *Nat. Chem. Biol.* **2020**, *16* (12), 1403–1410. https://doi.org/10.1038/s41589-020-0626-6.

(169) Ceccaldi, R.; Liu, J. C.; Amunugama, R.; Hajdu, I.; Primack, B.; Petalcorin, M. I. R.; O'Connor, K. W.; Konstantinopoulos, P. A.; Elledge, S. J.; Boulton, S. J.; Yusufzai, T.; D'Andrea, A. D. Homologous-Recombination-Deficient Tumours Are Dependent on

Polθ-Mediated Repair. *Nature* **2015**, *518* (7538), 258–262. https://doi.org/10.1038/nature14184.

(170) Çağlayan, M.; Wilson, S. In Vitro Assay to Measure DNA Polymerase β Nucleotide Insertion Coupled with the DNA Ligation Reaction during Base Excision Repair. *BIO-Protoc.* **2017**, *7* (12). https://doi.org/10.21769/BioProtoc.2341.

(171) Accomando, R. CHARACTERISTICS OF THE DNA POLYMERASE P-DIRECTED UNSCHEDULED DNA SYNTHESIS IN ISOLATED NUCLEI FROM ADULT RAT LIVER. 5.

(172) Steven, F. S.; Al-Habib, A. Inhibition of Trypsin and Chymotrypsin by Thiols. Biphasic Kinetics of Reactivation and Inhibition Induced by Sodium Periodate Addition. *Biochim. Biophys. Acta BBA - Enzymol.* **1979**, *568* (2), 408–415. https://doi.org/10.1016/0005-2744(79)90309-7.

(173) Tran, Q. T. N.; Tan, D. W. S.; Wong, W. S. F.; Chai, C. L. L. From Irreversible to Reversible Covalent Inhibitors: Harnessing the Andrographolide Scaffold for Anti-Inflammatory Action. *Eur. J. Med. Chem.* **2020**, *204*, 112481. https://doi.org/10.1016/j.ejmech.2020.112481.

(174) Pollard, D. A.; Pollard, T. D.; Pollard, K. S. Empowering Statistical Methods for Cellular and Molecular Biologists. *Mol. Biol. Cell* **2019**, *30* (12), 1359–1368. https://doi.org/10.1091/mbc.E15-02-0076.

(175) Sherman, M. S.; Lorenz, K.; Lanier, M. H.; Cohen, B. A. Cell-to-Cell Variability in the Propensity to Transcribe Explains Correlated Fluctuations in Gene Expression. *Cell Syst.* **2015**, *1* (5), 315–325. https://doi.org/10.1016/j.cels.2015.10.011.

(176) Hirose, F.; Hotta, Y.; Yamaguchi, M.; Matsukage, A. Difference in the Expression Level Od DNA Polymerase b in Mouse Tissues: High Expression in Pachytene Spermatocyte. *Exp. Cell Res.* **1989**, *181*, 169–180.

(177) Kidane, D.; Dalal, S.; Keh, A.; Liu, Y.; Zelterman, D.; Sweasy, J. B. DNA Polymerase Beta Is Critical for Genomic Stability of Sperm Cells. *DNA Repair* **2011**, *10* (4), 390–397. https://doi.org/10.1016/j.dnarep.2011.01.003.

(178) Calafiore, M.; Copani, A.; Deng, W. DNA Polymerase-β Mediates the Neurogenic Effect of β-Amyloid Protein in Cultured Subventricular Zone Neurospheres. *J. Neurosci. Res.* **2012**, *90* (3), 559–567. https://doi.org/10.1002/jnr.22780.

(179) Ray, S.; Breuer, G.; DeVeaux, M.; Zelterman, D.; Bindra, R.; Sweasy, J. B. DNA Polymerase Beta Participates in DNA End-Joining. *Nucleic Acids Res.* **2018**, *46* (1), 242–255. https://doi.org/10.1093/nar/gkx1147.

(180) Abbotts, J.; SenGupta, D. N.; Zmudzka, B.; Widen, S. G.; Notario, V.; Wilson, S. H. Expression of Human DNA Polymerase .Beta. in Escherichia Coli and Characterization of the Recombinant Enzyme. *Biochemistry* **1988**, *27* (3), 901–909. https://doi.org/10.1021/bi00403a010.

(181) Prasad, R.; Kumar, A.; Widen, S.; Casas-Finet, J.; Wilson, S. Identification of Residues in the Single-Stranded DNA-Binding Site of the 8-KDa Domain of Rate DNA Polymerase b by UV Cross-Linking. *J. Biol. Chem.* **1993**, *268*, 22746–22755.

(182) Beard, W.; Wilson, S. Purification and Domain-Mapping of Mammalian DNA Polymerase b. *Methods Enzymol.* **1995**, *262*, 98–107.

(183) Garcıa-Dıaz, M.; Bebenek, K.; Sabariegos, R.; Domınguez, O.; Rodrıguez, J.; Kirchhoff, T.; Garcıa-Palomero, E.; Picher, A. J. DNA Polymerase λ, a Novel DNA Repair Enzyme in Human Cells. *J. Biol. Chem.* **2002**, *277*, 13184–13191.

(184) Fiala, K. A.; Abdel-Gawad, W.; Suo, Z. Pre-Steady-State Kinetic Studies of the Fidelity and Mechanism of Polymerization Catalyzed by Truncated Human DNA Polymerase λ †. *Biochemistry* **2004**, *43* (21), 6751–6762. https://doi.org/10.1021/bi049975c.

(185) Xu, P.; Chen, D.; Xi, J.; Yao, Z. Short Protecting Group-free Syntheses of Camptothecin and 10-Hydroxycamptothecin Using Cascade Methodologies. *Chem. - Asian J.* **2015**, *10* (4), 976–981.

(186) McGuigan, C.; Serpi, M.; Slusarczyk, M.; Ferrari, V.; Pertusati, F.; Meneghesso, S.; Derudas, M.; Farleigh, L.; Zanetta, P.; Bugert, J. Anti-flavivirus Activity of Different Tritylated Pyrimidine and Purine Nucleoside Analogues. **2016**, 9.

(187) Sheng, J.; Hassan, A. E. A.; Zhang, W.; Zhou, J.; Xu, B.; Soares, A. S.; Huang, Z. Synthesis, Structure and Imaging of Oligodeoxyribonucleotides with Tellurium-Nucleobase Derivatization. *Nucleic Acids Res.* **2011**, *39* (9), 3962–3971. https://doi.org/10.1093/nar/gkq1288.

(188) Karino, N. Synthesis and Properties of Oligonucleotides Containing 5-Formyl-2'-Deoxycytidine: In Vitro DNA Polymerase Reactions on DNA Templates Containing 5-Formyl-2'-Deoxycytidine. *Nucleic Acids Res.* **2001**, *29* (12), 2456–2463. https://doi.org/10.1093/nar/29.12.2456.

(189) Neef, A.; Luedtke, N. An Azide-Modified Nucleoside for Metabolic Labeling of DNA. *ChemBioChem 15* (6), 789–793.

(190) Barvian, M. R.; Greenberg, M. M. Independent Generation of the Major Addnet of Hydroxyl Radical and Thymidine. Examination of Intramoleenlar Hydrogen Atom Transfer in Competition With Thiol Trapping. *Tetrahedron Lett.* **1992**, *33* (41), 6057–6060.

(191) Khalil, A.; Bayach, I.; Mathé, C. A Synergistic Synthetic and Computational Insights towards Anomerization of N-Nitro Pyrimidine Nucleosides Using Fluorinating Agents. *J. Fluor. Chem.* **2020**, *233*, 109504. https://doi.org/10.1016/j.jfluchem.2020.109504.

(192) Fan, H.; Sun, H.; Haque, M. M.; Peng, X. Effect of Triazole-Modified Thymidines on DNA and RNA Duplex Stability. *ACS Omega* **2019**, *4* (3), 5107–5116. https://doi.org/10.1021/acsomega.9b00089.

(193) Wen, Z.; Peng, J.; Tuttle, P. R.; Ren, Y.; Garcia, C.; Debnath, D.; Rishi, S.; Hanson, C.; Ward, S.; Kumar, A.; Liu, Y.; Zhao, W.; Glazer, P. M.; Liu, Y.; Sevilla, M. D.; Adhikary, A.; Wnuk, S. F. Electron-Mediated Aminyl and Iminyl Radicals from C5 Azido-Modified Pyrimidine Nucleosides Augment Radiation Damage to Cancer Cells. *Org Lett* **2018**, 5.

(194) Miyata, K.; Tamamushi, R.; Ohkubo, A.; Taguchi, H.; Seio, K.; Santa, T.; Sekine, M. Synthesis and Properties of a New Fluorescent Bicyclic 4-N-Carbamoyldeoxycytidine Derivative. *Org Lett* **2006**, *8* (8), 4.

(195) Li, T.; Rokita, S. Selective Modification of DNA Controlled by an Ionic Signal. *J Am Chem Soc* **1991**, *113* (20), 7771–7773.

(196) Handlon, A.; Oppenheimer, N. Thiol Reduction of 3'-Azidothymidine to 3'-Aminothymidine: Kinetics and Biomedical Implications. *Pharm Res* **1998**, *5*, 297–299.

(197) Vanheusden, V.; Munier-Lehmann, H.; Froeyen, M.; Dugue, L.; Heyerick, A.; Keukeleire, D. D.; Pochet, S.; Busson, R.; Herdewijn, P.; Calenbergh, S. V. 3′-C-Branched-Chain-

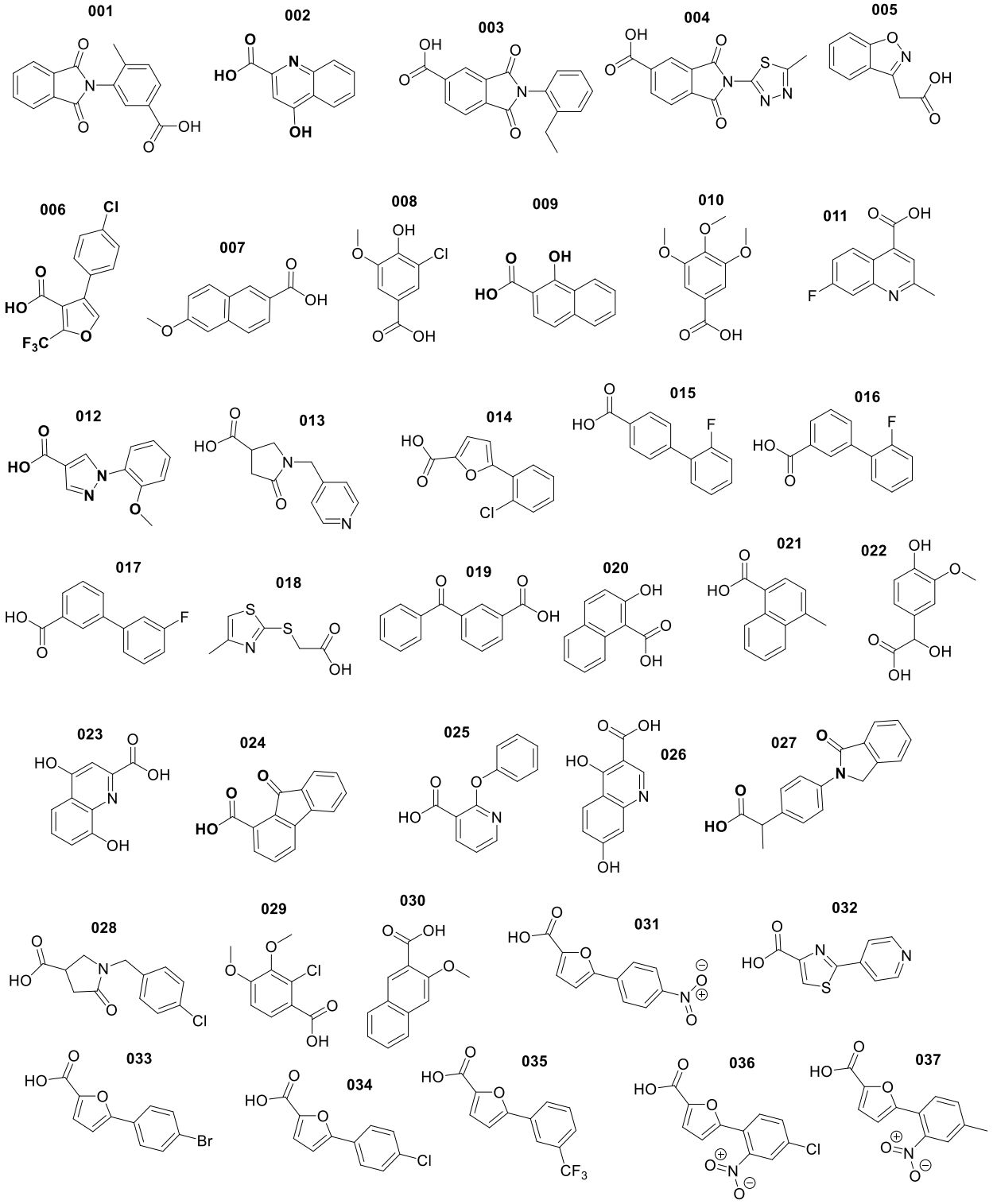
Substituted Nucleosides and Nucleotides as Potent Inhibitors of Mycobacterium Tuberculosis Thymidine Monophosphate Kinase. 11.

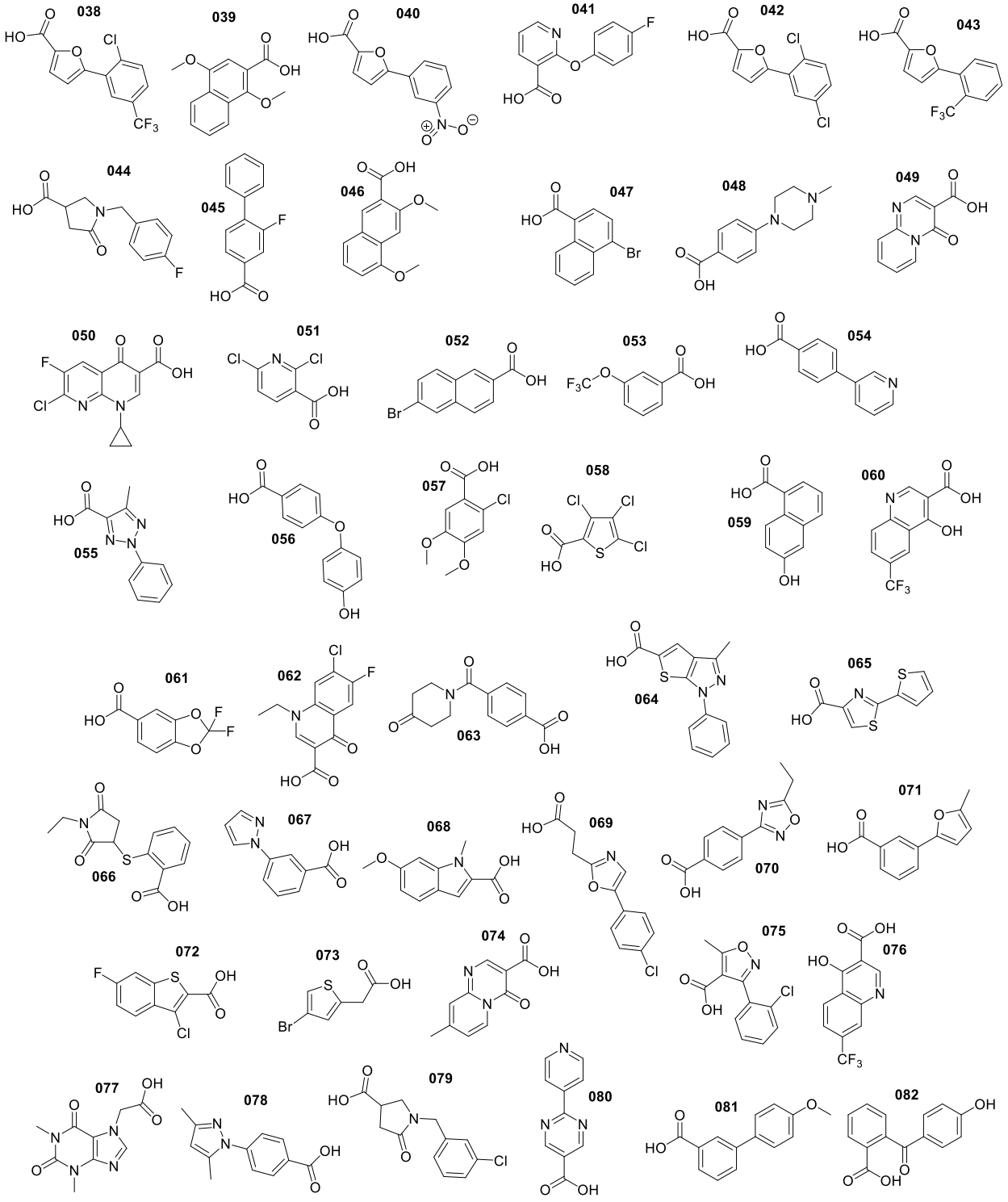
(198) Vernekar, S. K. V.; Qiu, L.; Zhang, J.; Kankanala, J.; Li, H.; Geraghty, R. J.; Wang, Z. 5′-Silylated 3′-1,2,3-Triazolyl Thymidine Analogues as Inhibitors of West Nile Virus and Dengue Virus. *J. Med. Chem.* **2015**, *58* (9), 4016–4028. https://doi.org/10.1021/acs.jmedchem.5b00327.

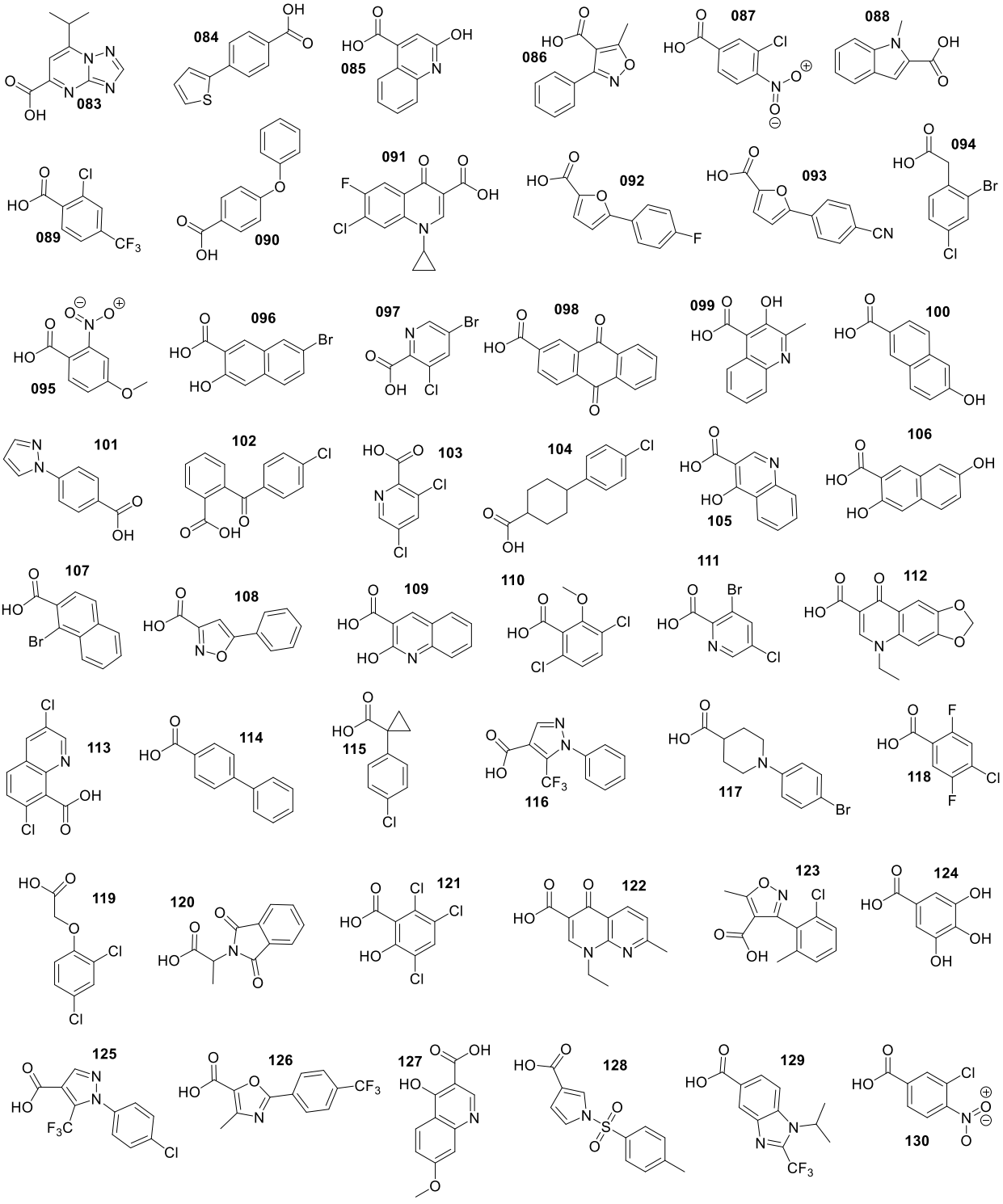
(199) Ikawa, H.; Nishimura, M. Preparation of Aromatic Amides Having Hydroxy Group in the Ortho Position. JP 08295656.

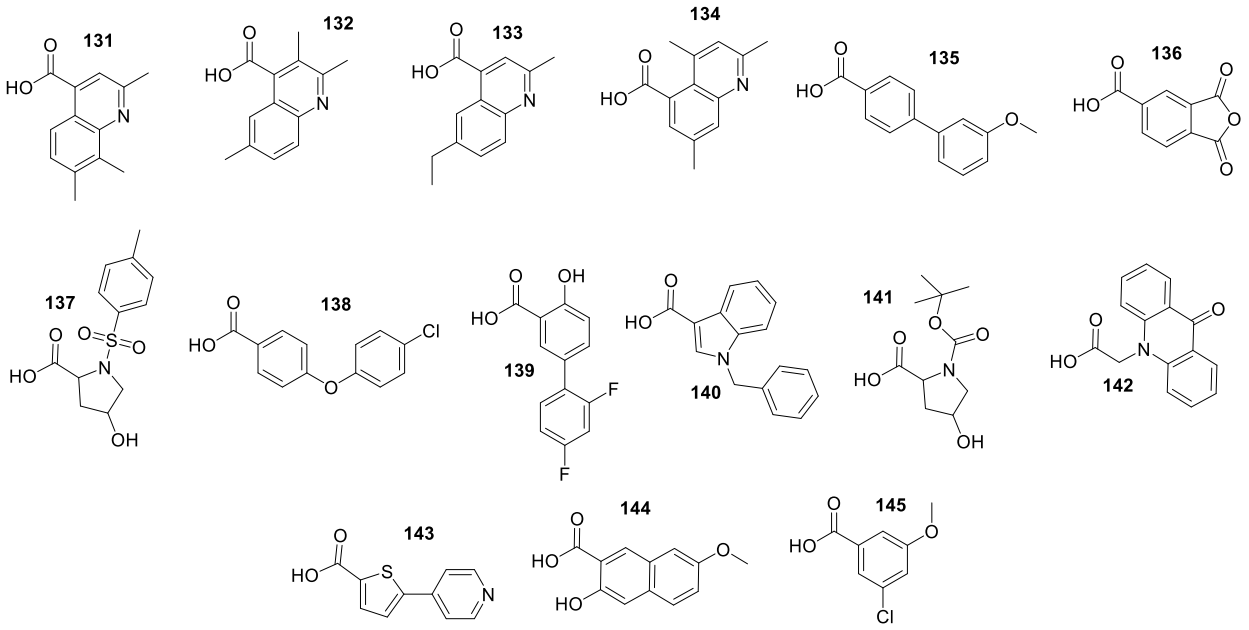
Appendices

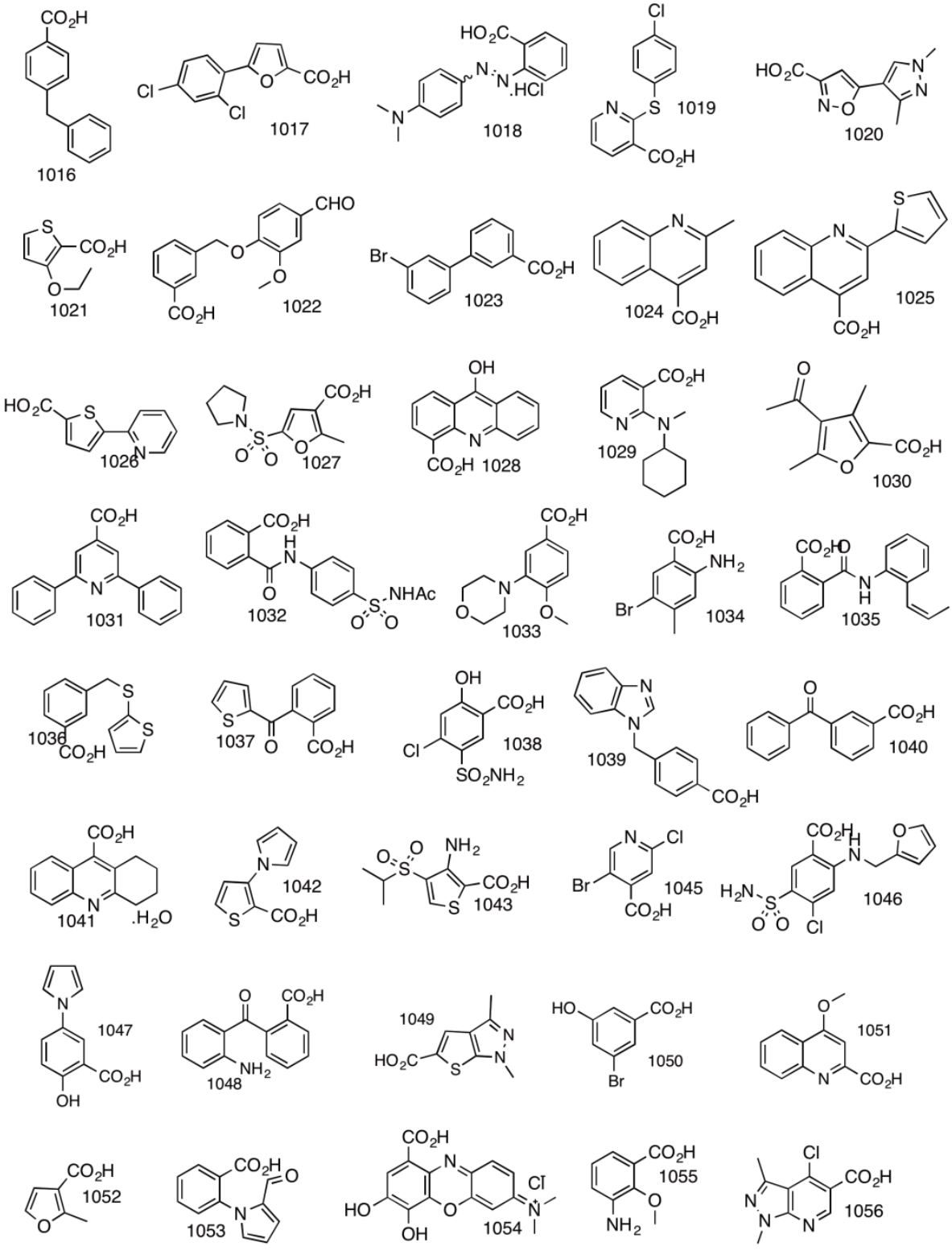
Appendix 1. Carboxylic acids used to prepare inhibitor candidates.

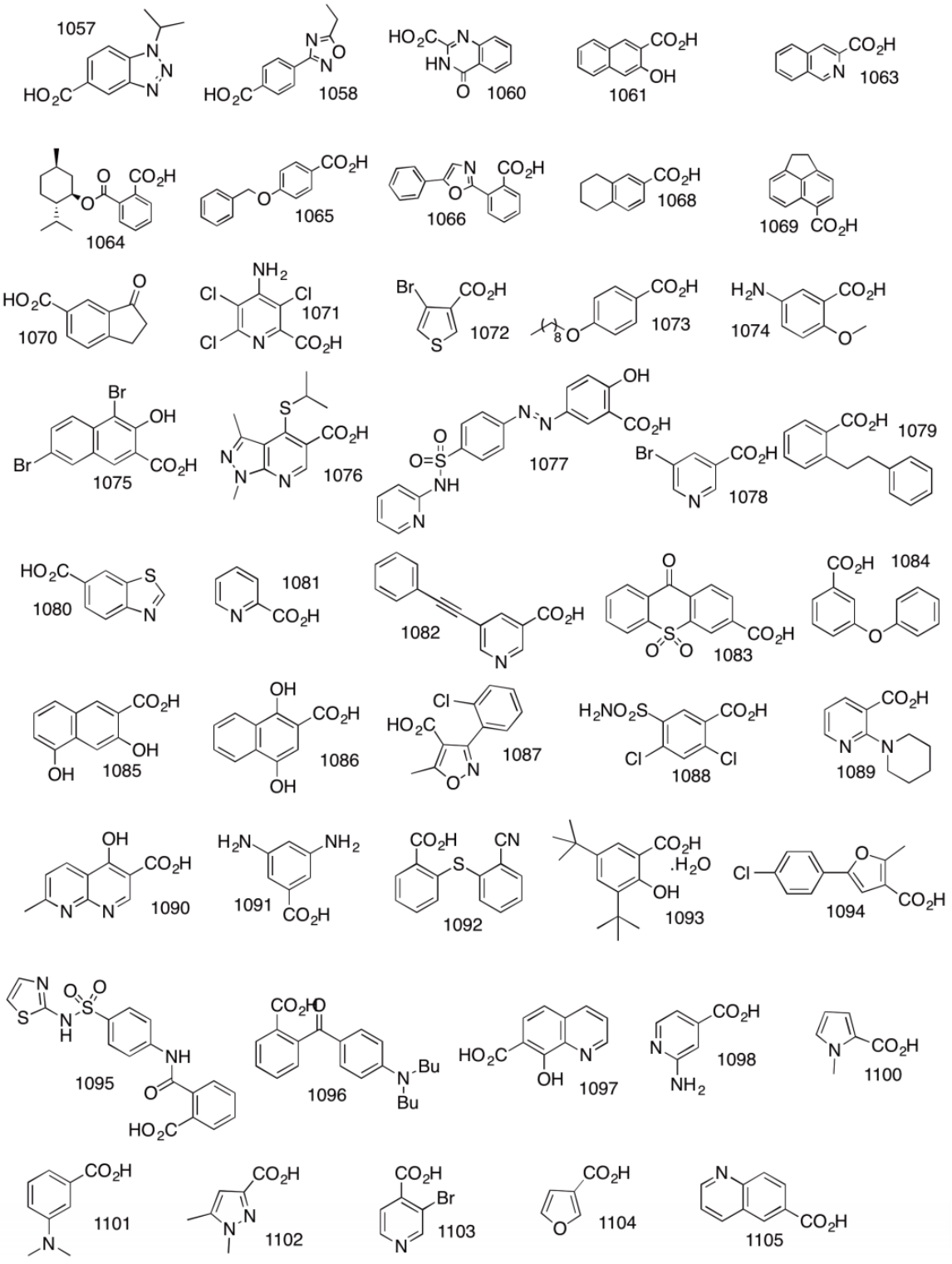


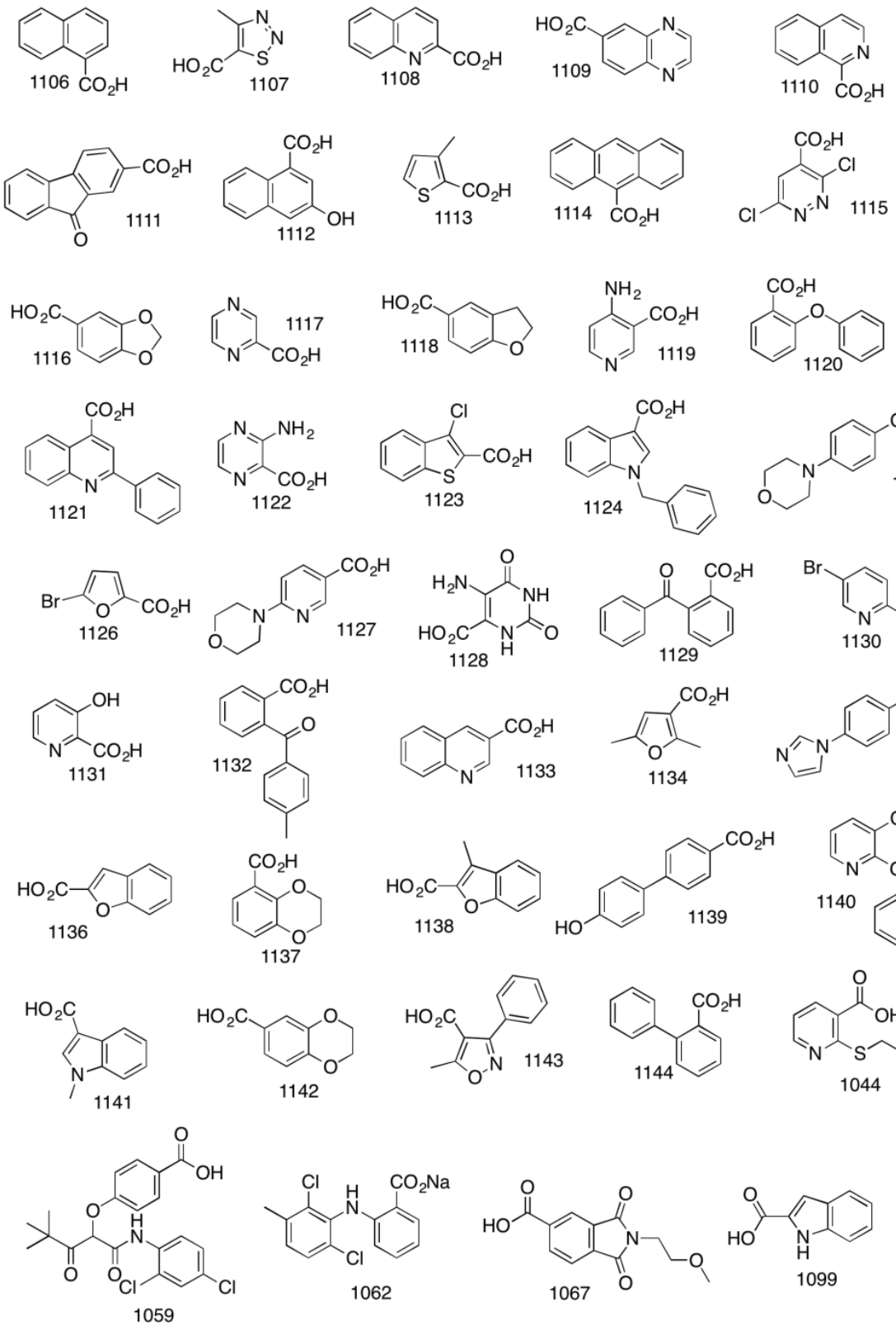

038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053
054
055
056
057
058
059
060
061
062
063
064
065
066
067
068
069
070
071
072
073
074
075
076
077
078
079
080
081
082

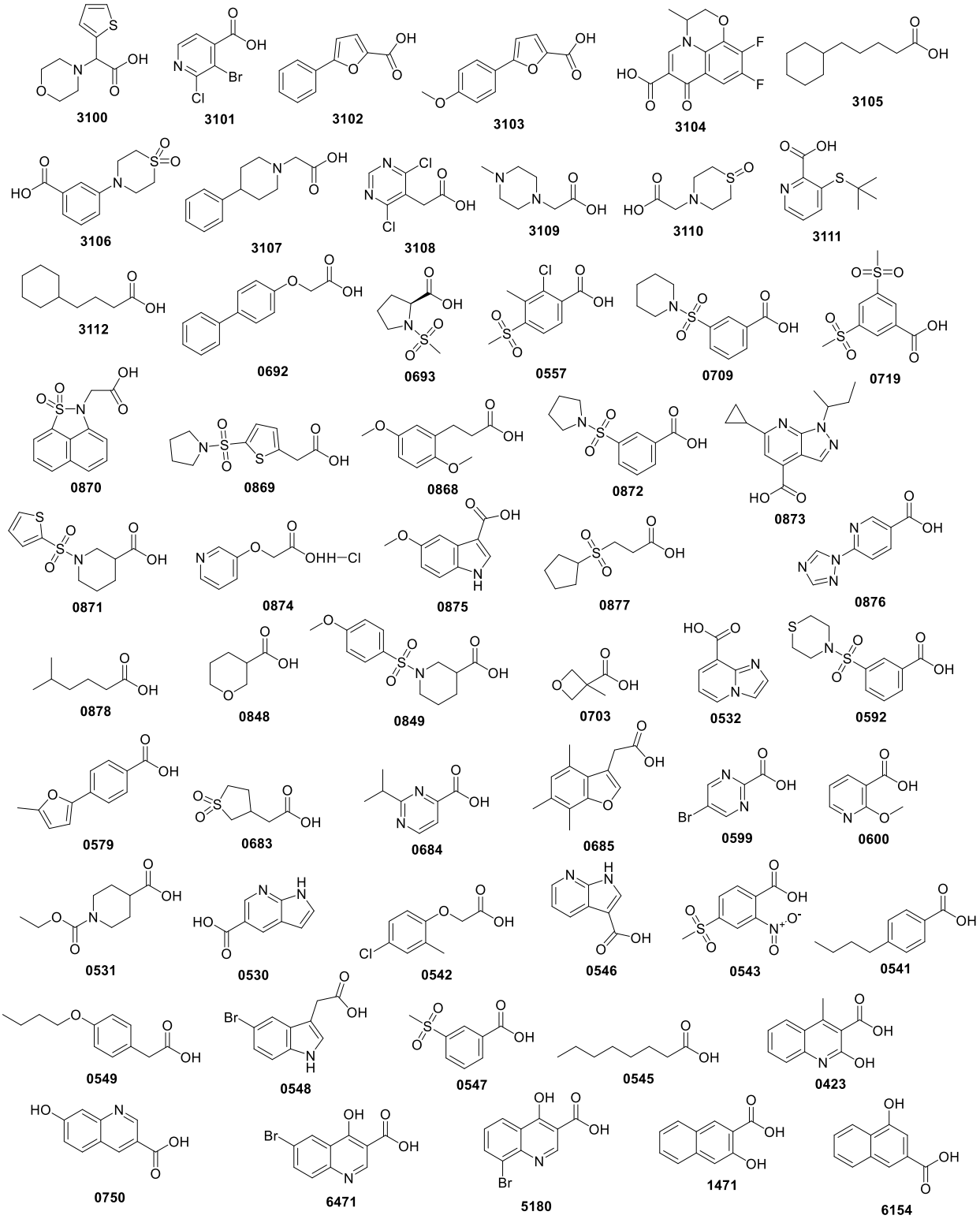

083
084
085
086
087
088
089
090
091
092
093
094
095
096
097
098
099
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130

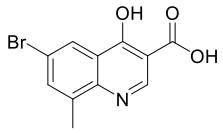

131
132
133
134
135
136
137
138
139
140
141
142
143
144
145


1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056

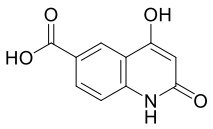

1057
1058
1060
1061
1063
1064
1065
1066
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1100
1101
1102
1103
1104
1105


1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1044
1059
1062
1067
1099

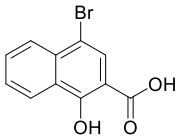

3100
3101
3102
3103
3104
3105
3106
3107
3108
3109
3110
3111
3112
0692
0693
0557
0709
0719
0870
0869
0868
0872
0873
0871
0874
0875
0877
0876
0878
0848
0849
0703
0532
0592
0579
0683
0684
0685
0599
0600
0531
0530
0542
0546
0543
0541
0549
0548
0547
0545
0423
0750
6471
5180
1471
6154



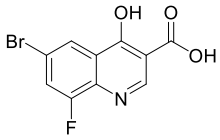
6105



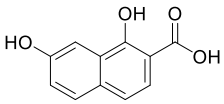
7405



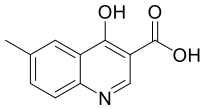
5653



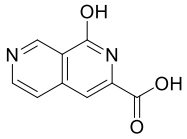
4420



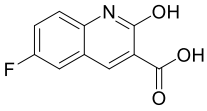
0867



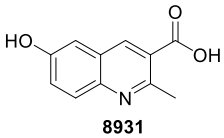
7146



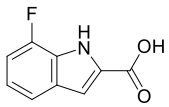
1834



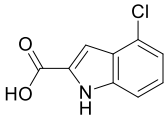
9658



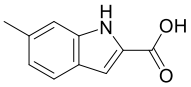
8931



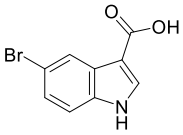
9666



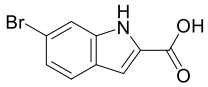
0151



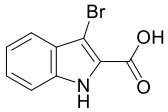
0152



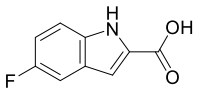
0153



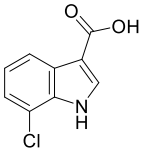
9665



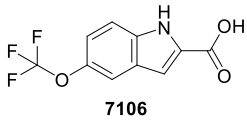
0155



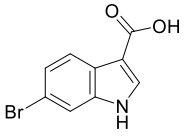
9664



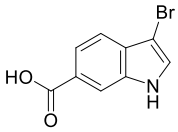
7105



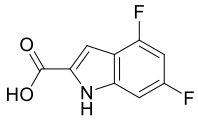
7106



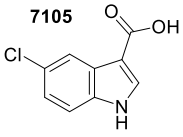
4353



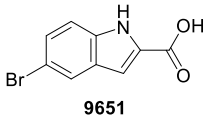
9748



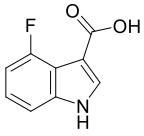
4350



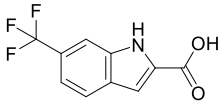
9650



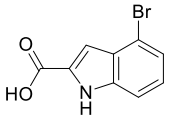
9651



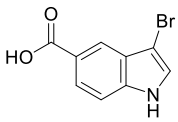
0150



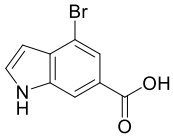
0149



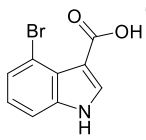
4351



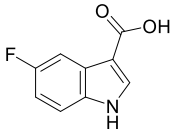
7107



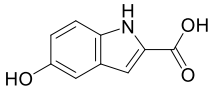
4352



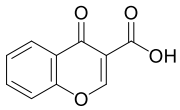
0170



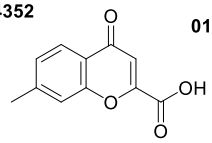
4349



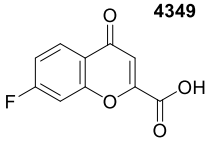
0154



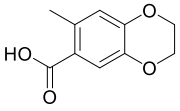
7109



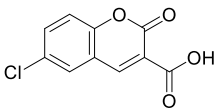
9751



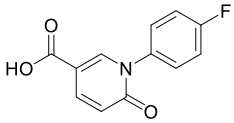
9668



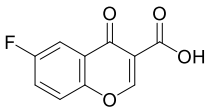
4304



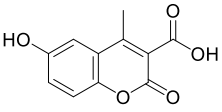
5928



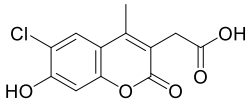
5597



3775



5596



4869

4870

5593

9752

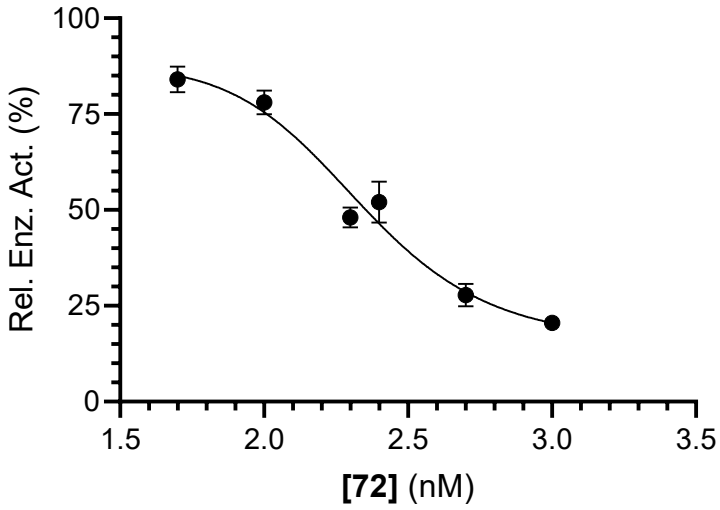
5598

5594

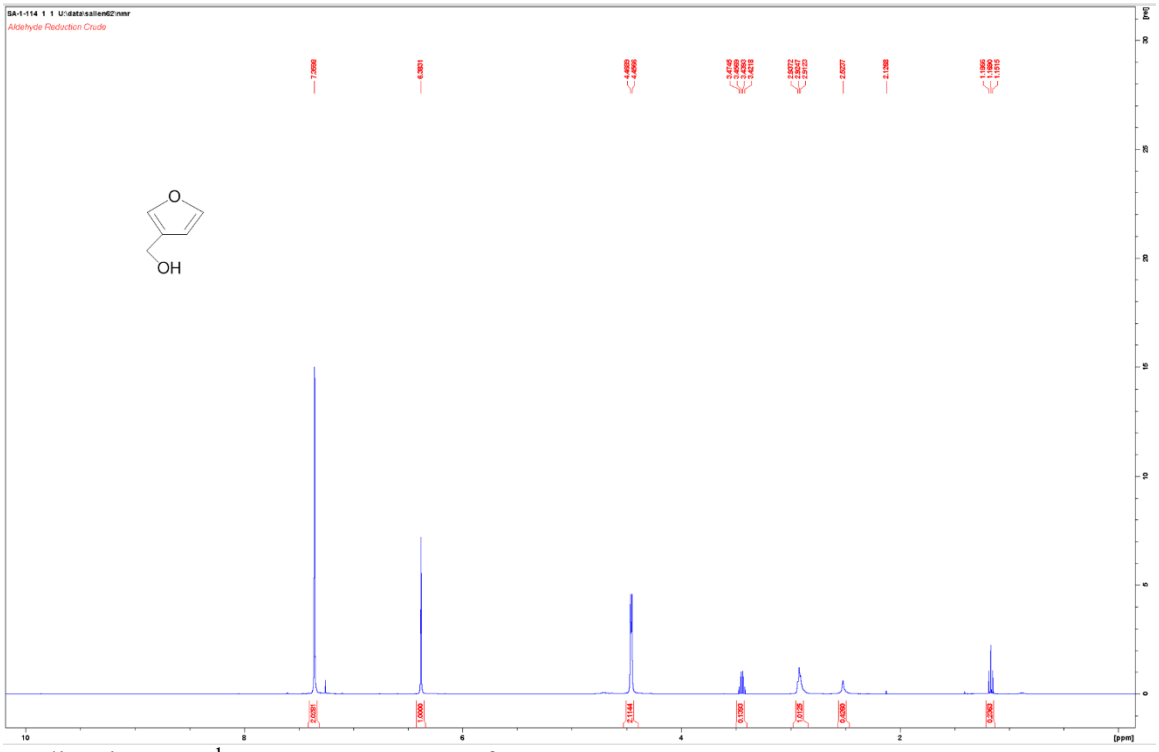
5595

7108

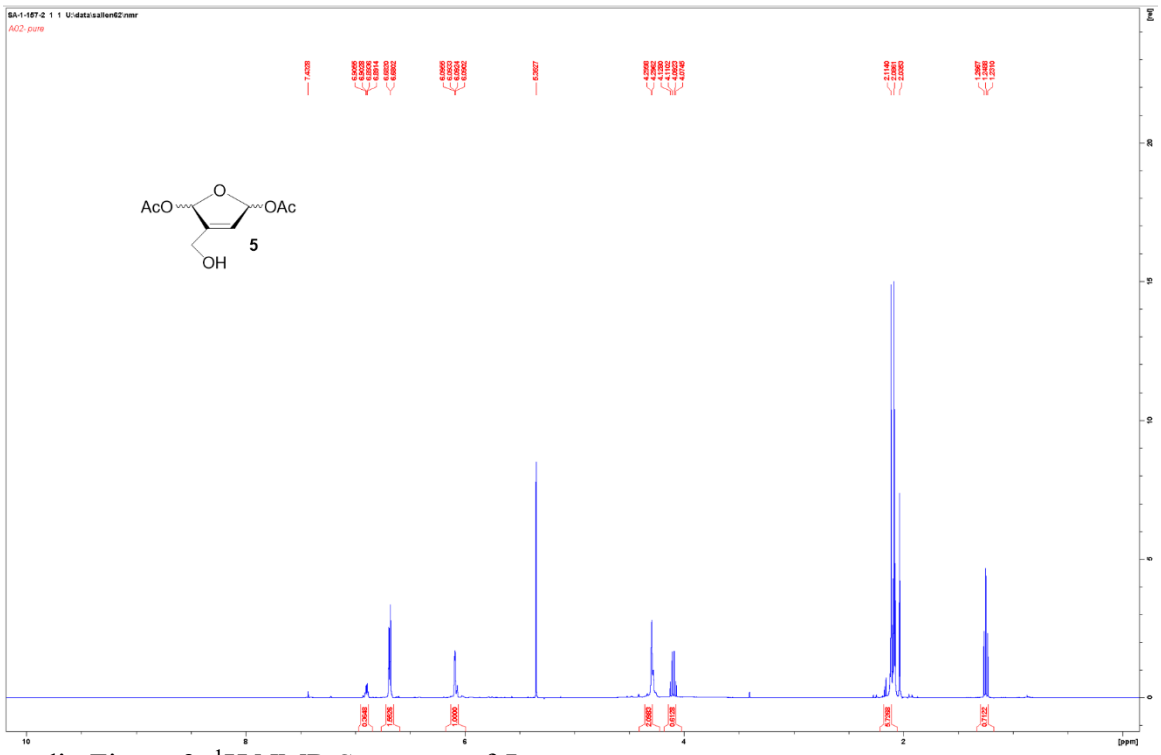
Appendix 2. Supplementary Data including NMR Spectra and MALDI-TOF



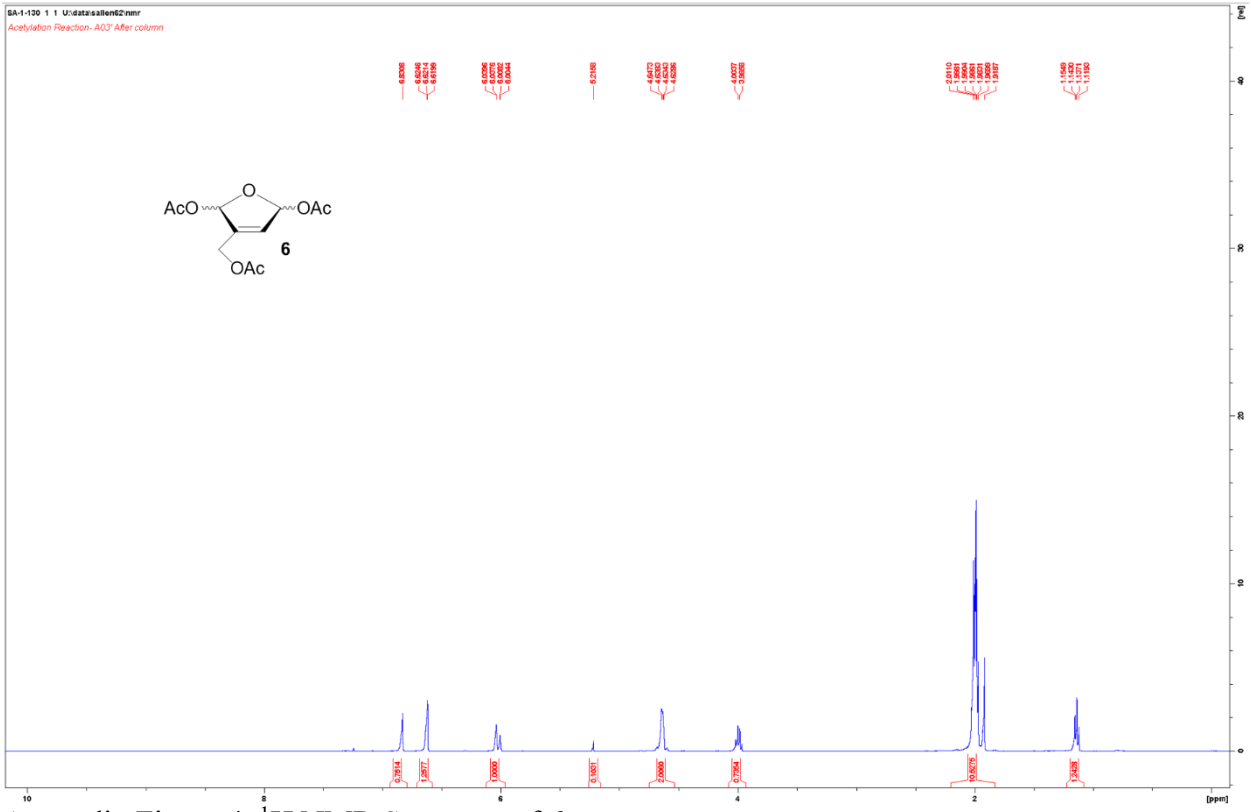
Appendix Figure 1. IC_{50} of **72** under other conditions.



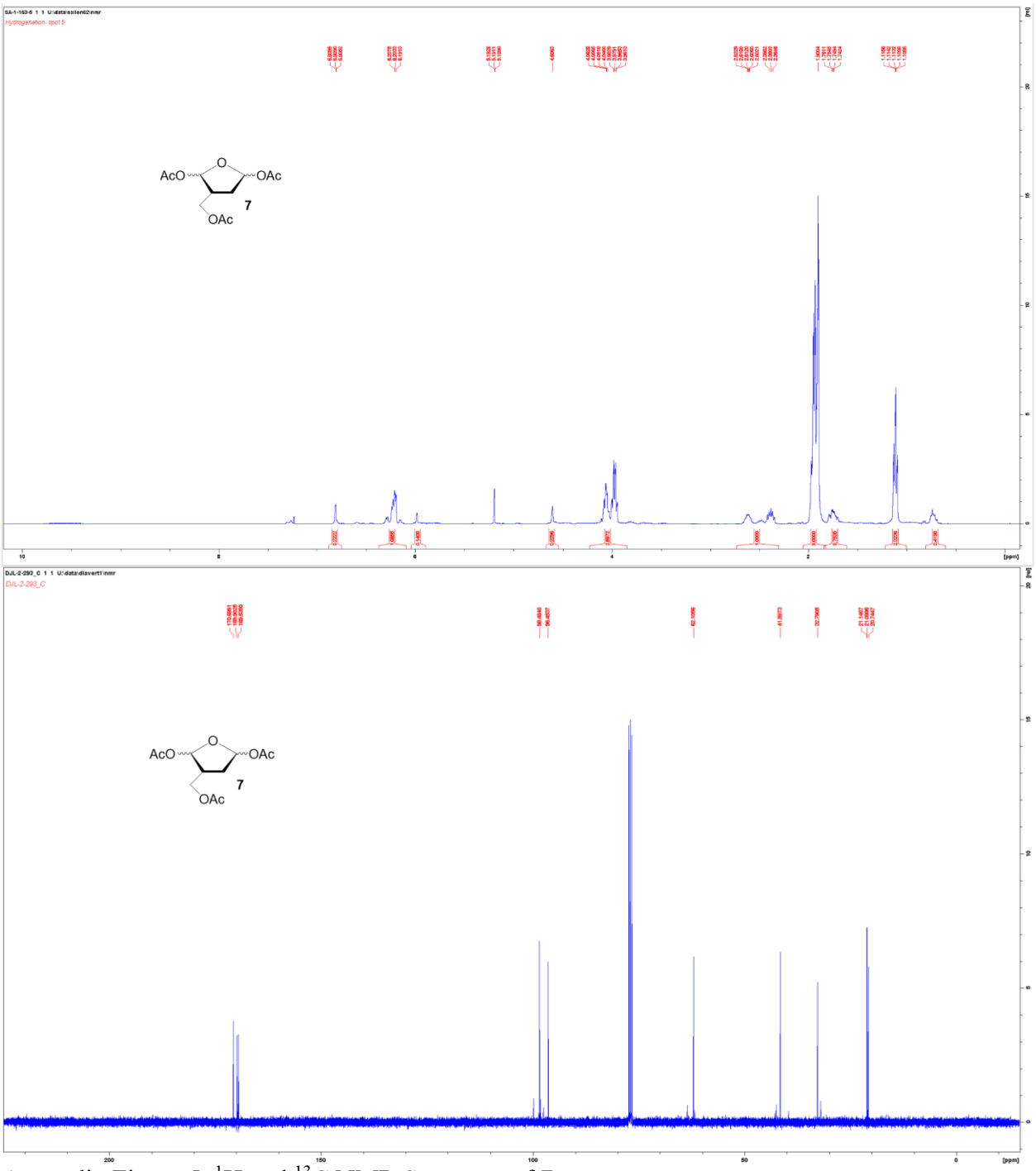
Appendix Figure 2. ^{1}H NMR Spectrum of precursor to **5**.



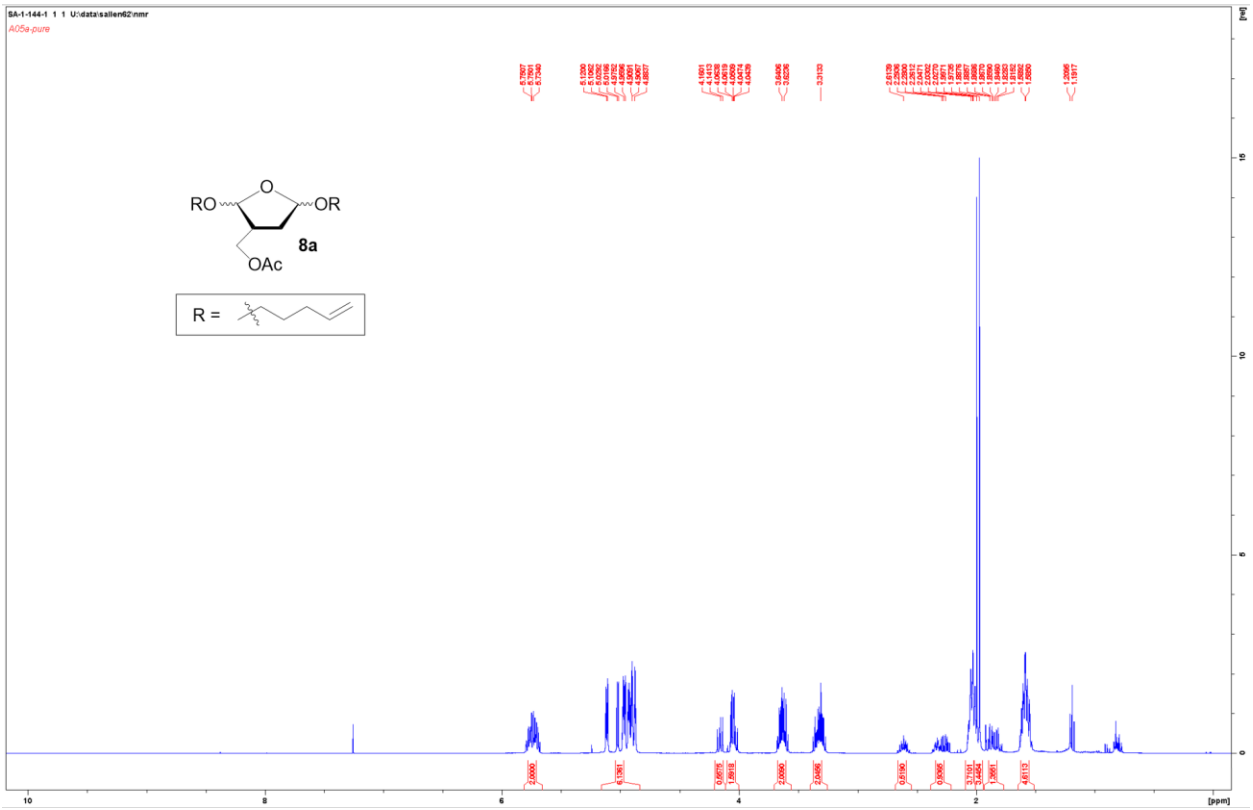
Appendix Figure 3. ^{1}H NMR Spectrum of **5**.



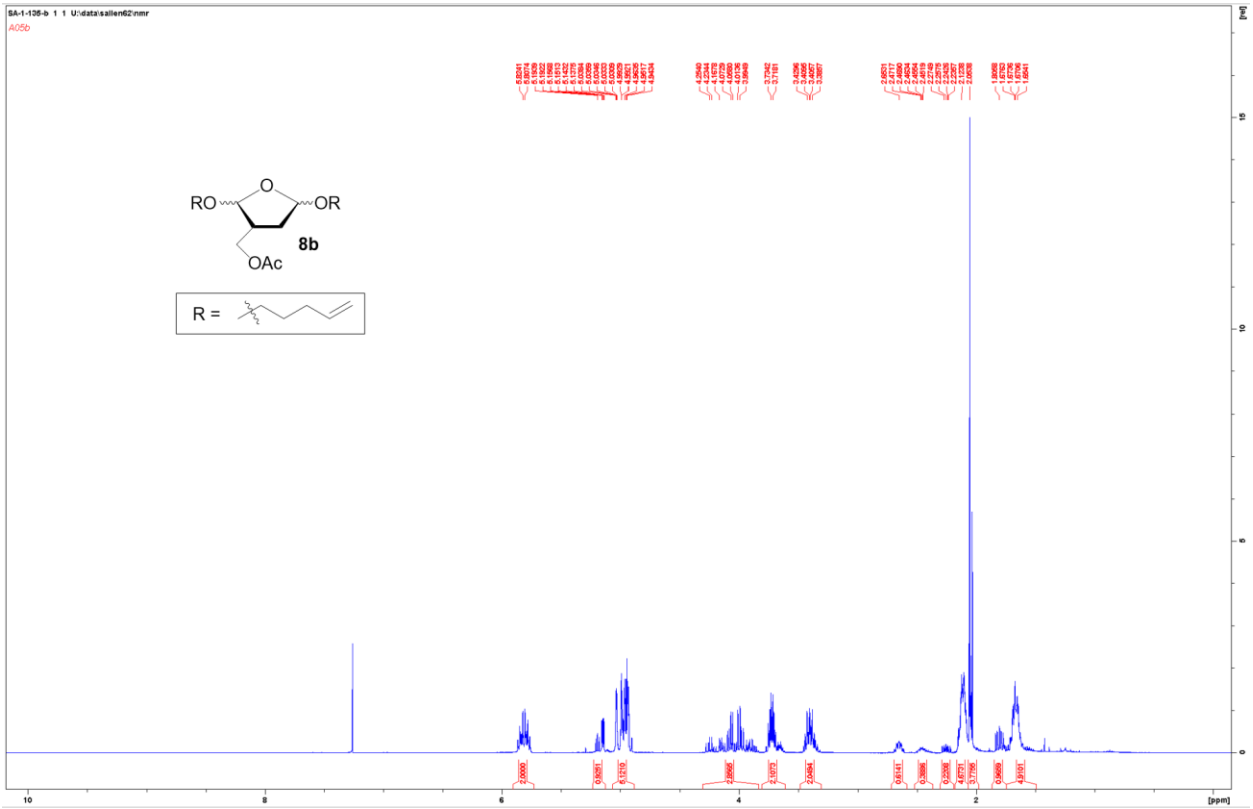
Appendix Figure 4. ^{1}H NMR Spectrum of **6**.



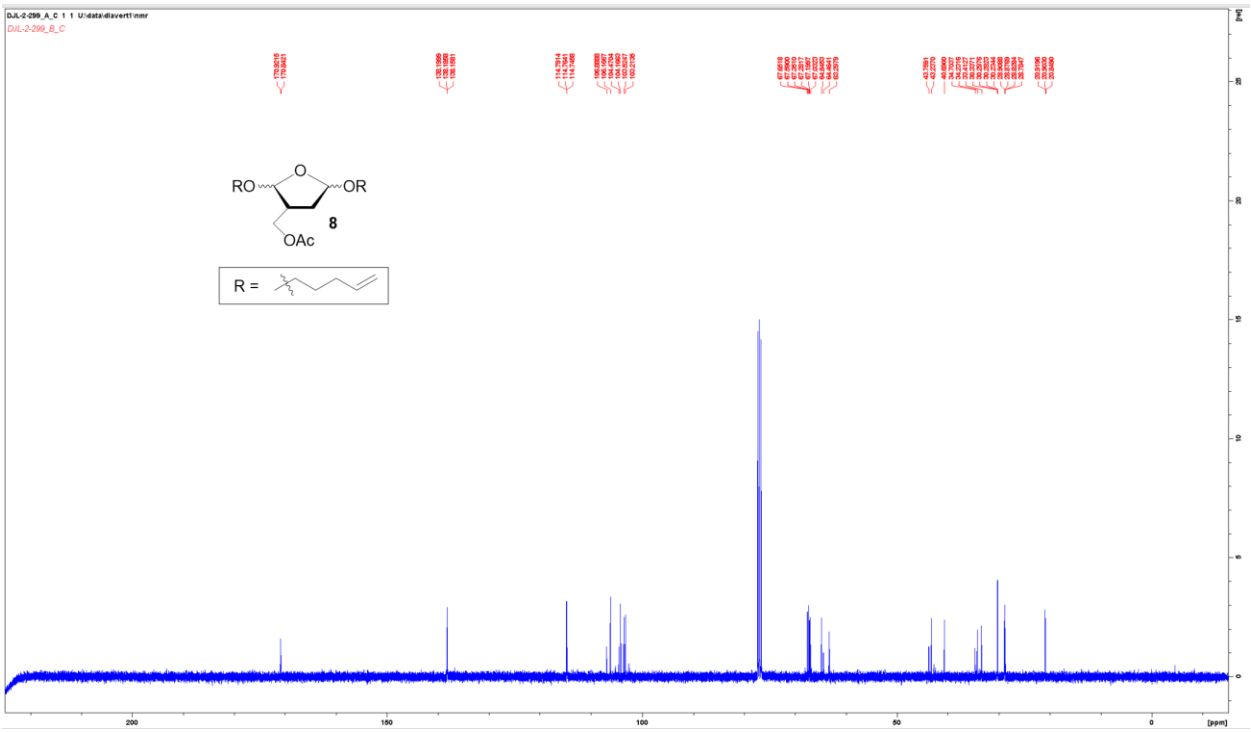
Appendix Figure 5. ^{1}H and ^{13}C NMR Spectrum of **7**.



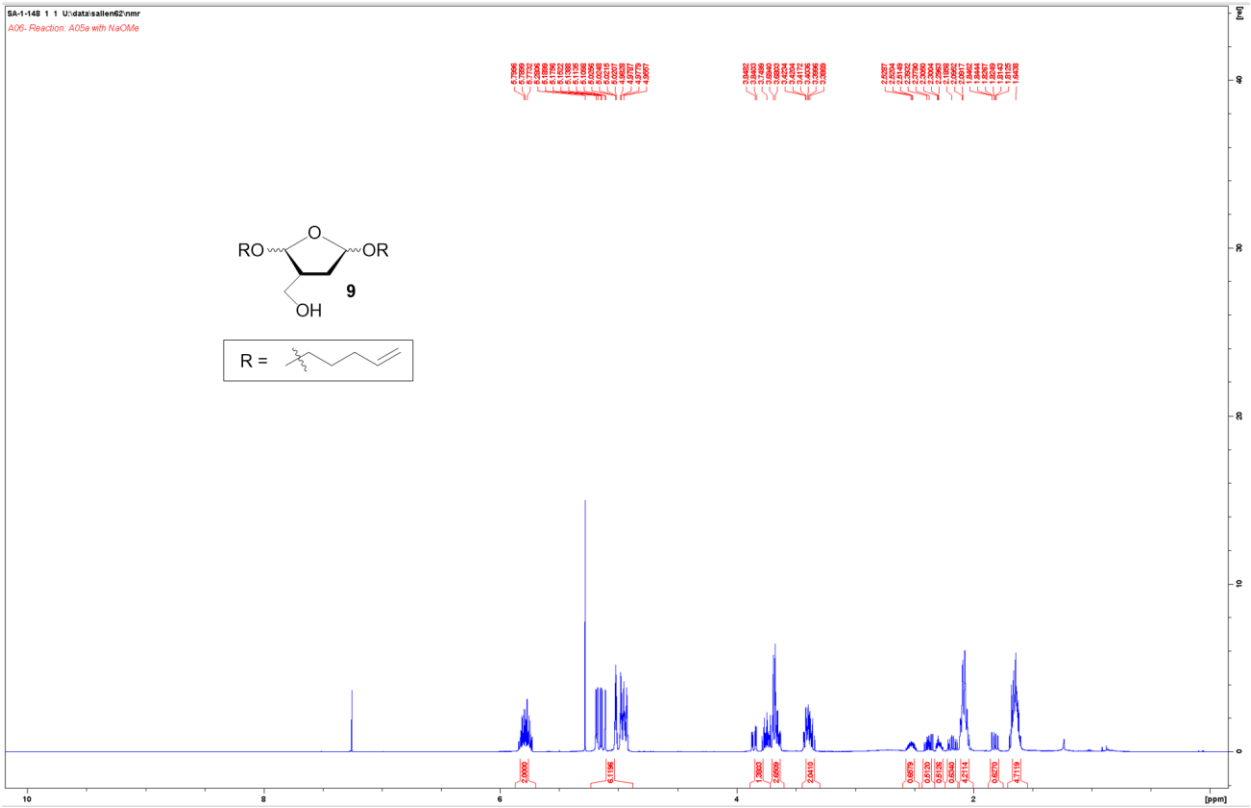
Appendix Figure 6. ^{1}H NMR Spectrum of **8a**.



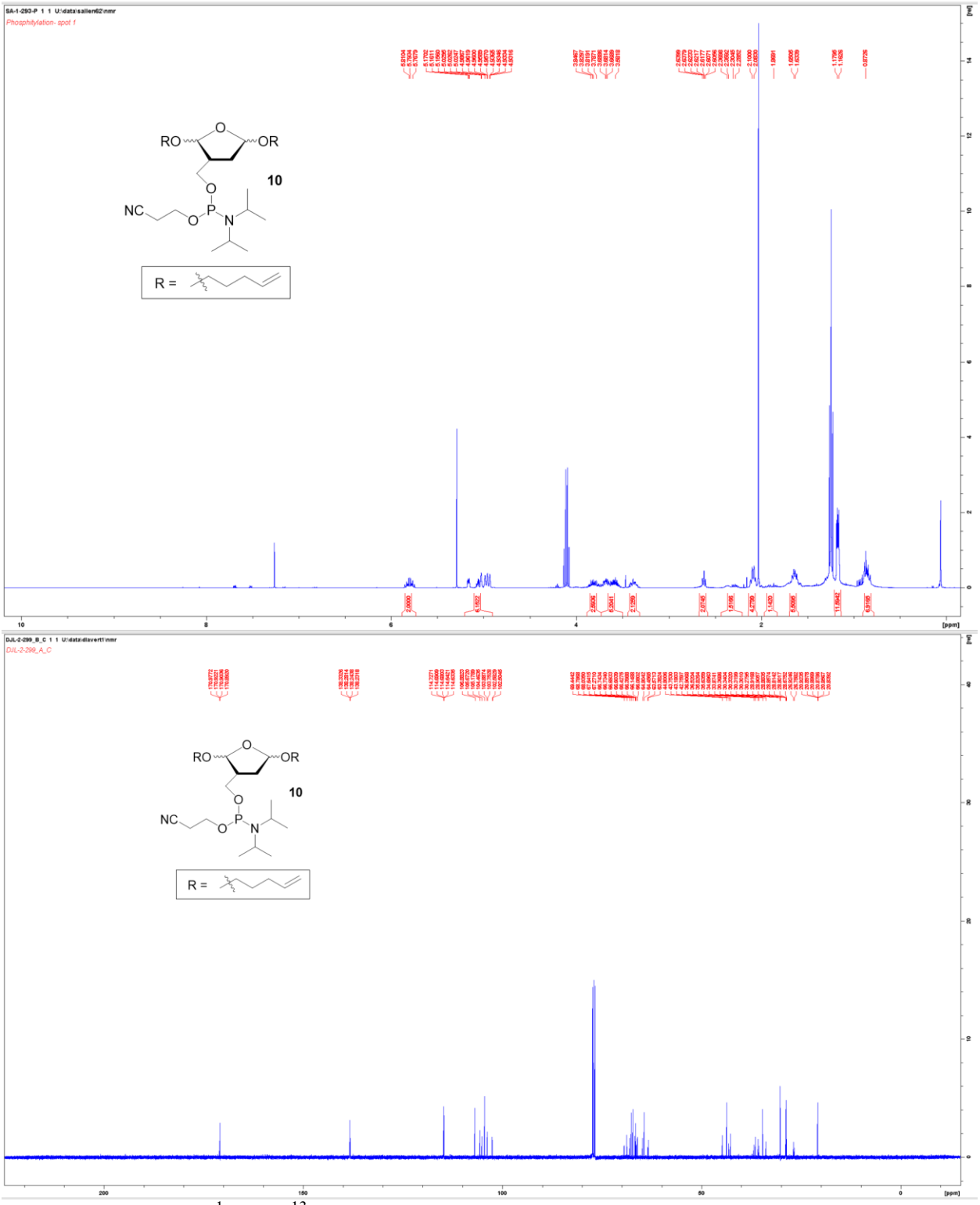
Appendix Figure 7. ^{1}H NMR Spectrum of **8b**.



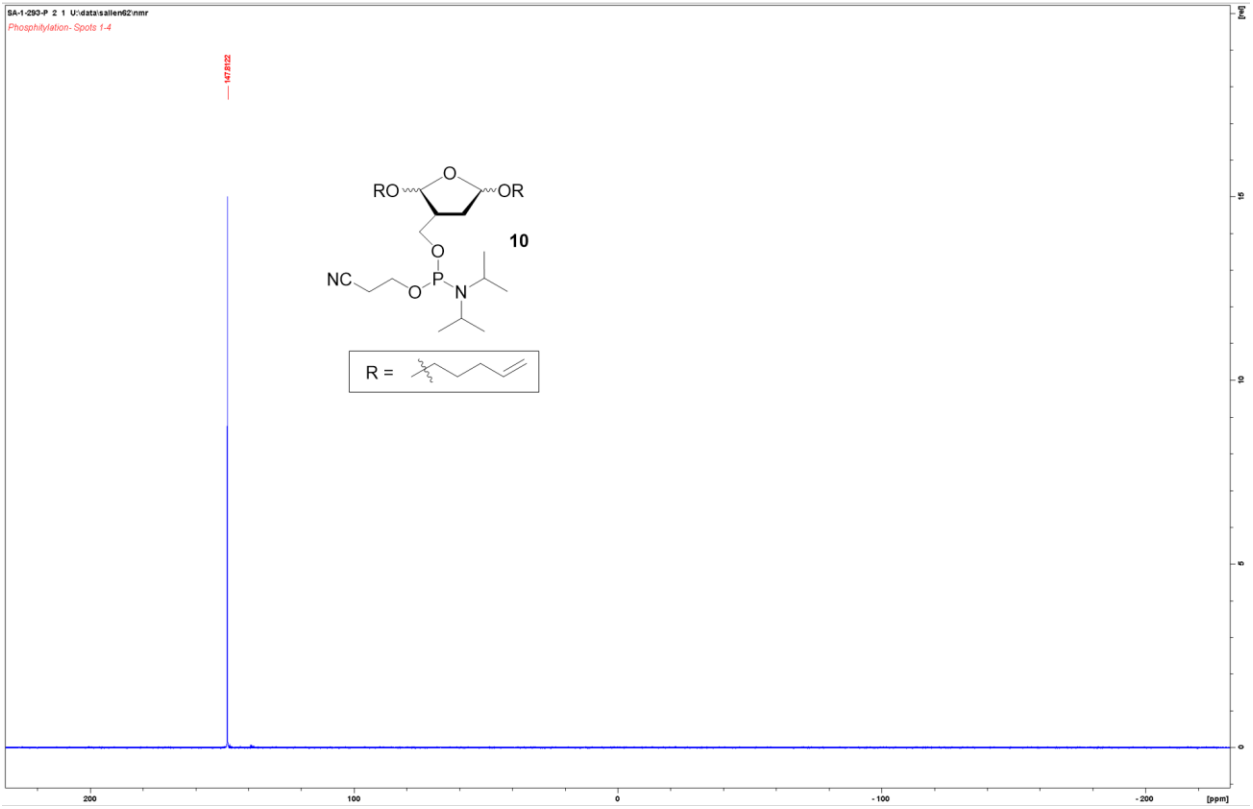
Appendix Figure 8. ^{13}C NMR Spectrum of **8a** and **8b**.



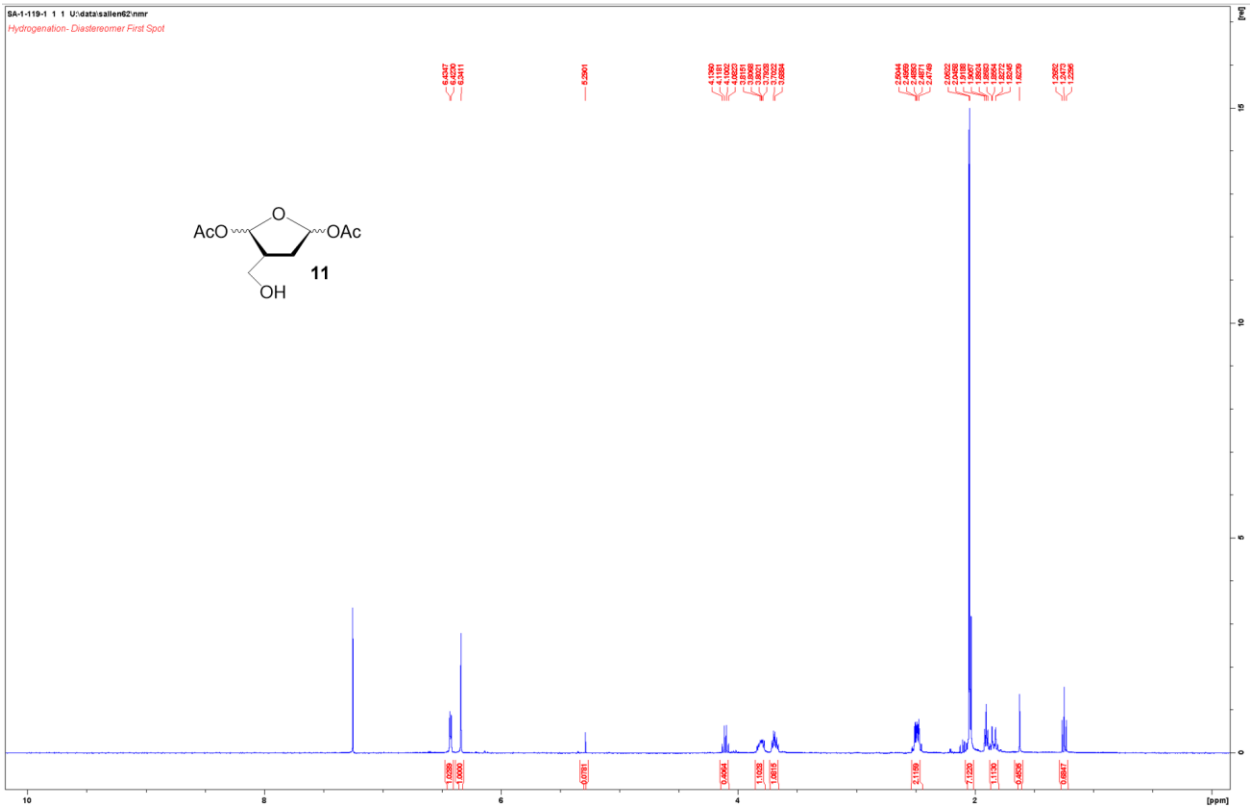
Appendix Figure 9. ^{1}H NMR Spectrum of **9**.



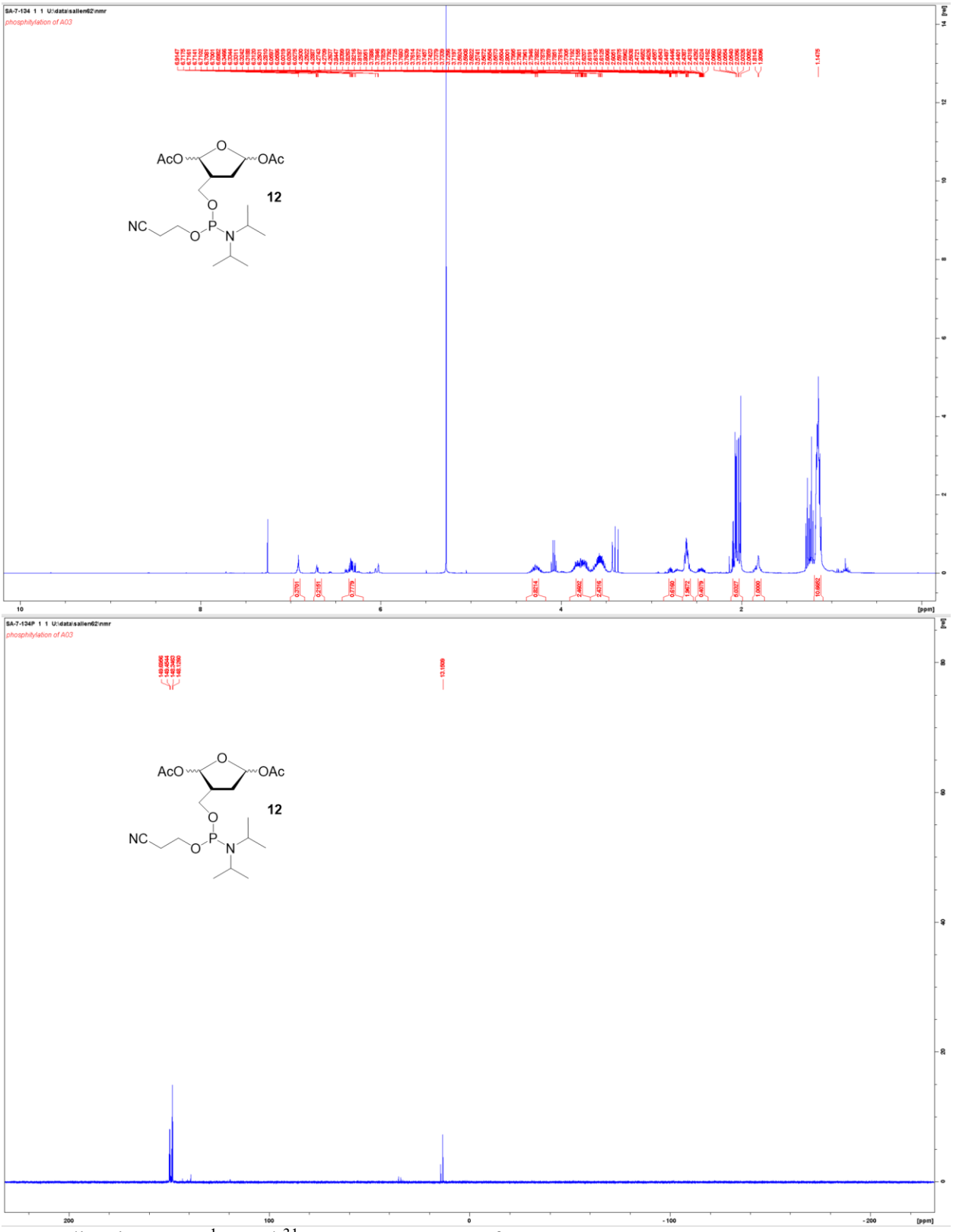
Appendix Figure 10. ^{1}H and ^{13}C NMR Spectrum of **10**.



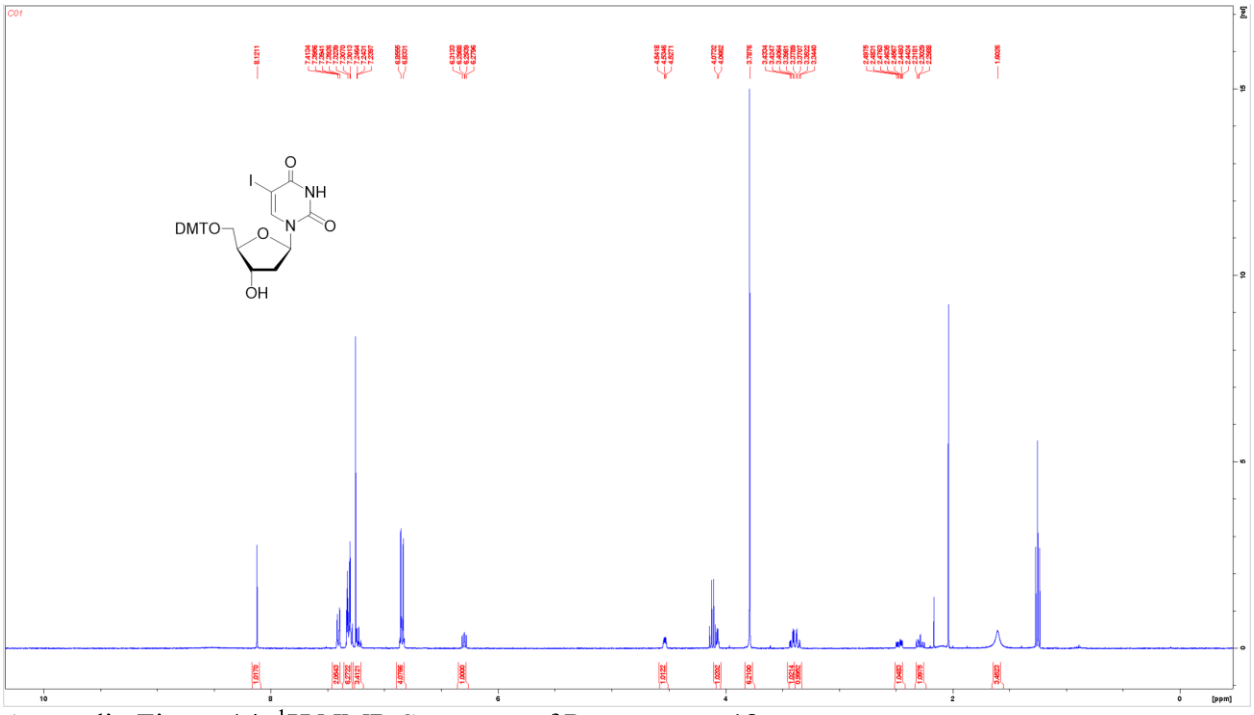
Appendix Figure 11. ^{31}P NMR Spectrum of **10**.



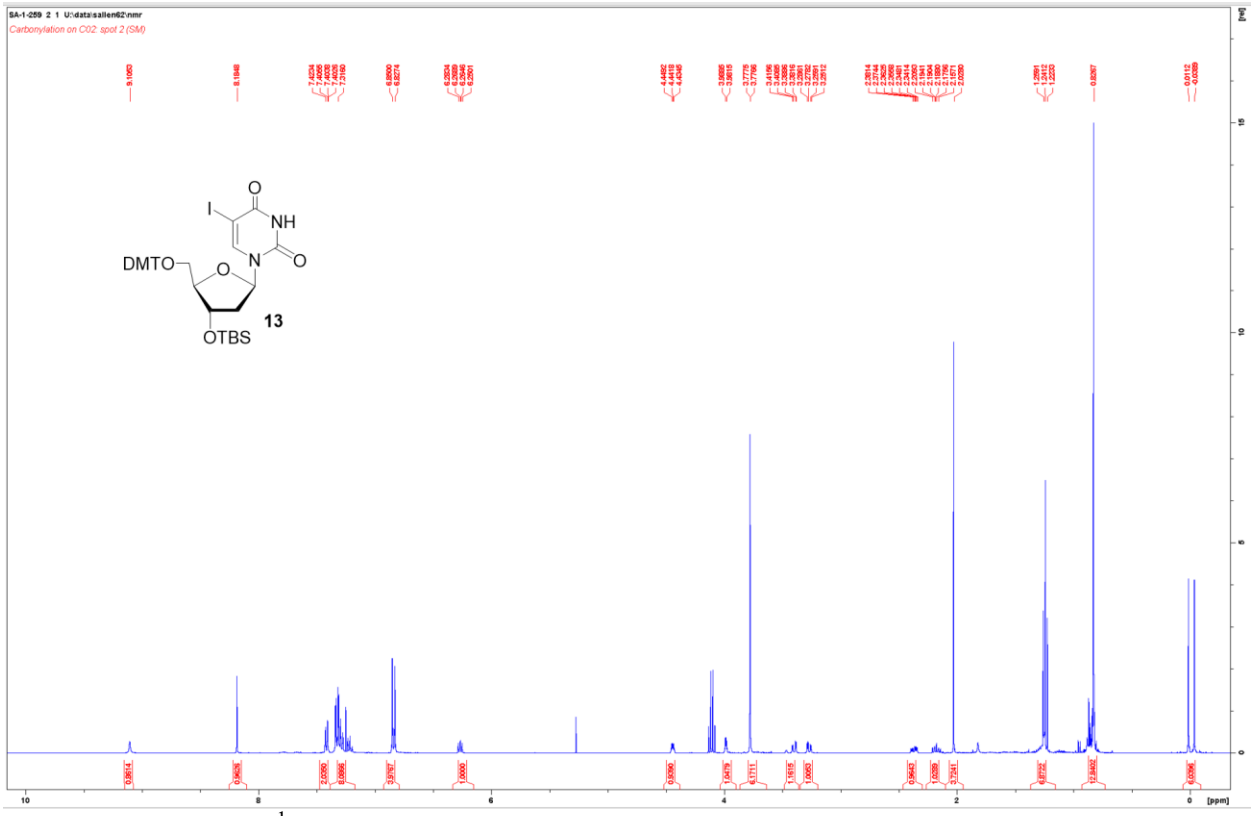
Appendix Figure 12. ^{1}H NMR Spectrum of **11**.



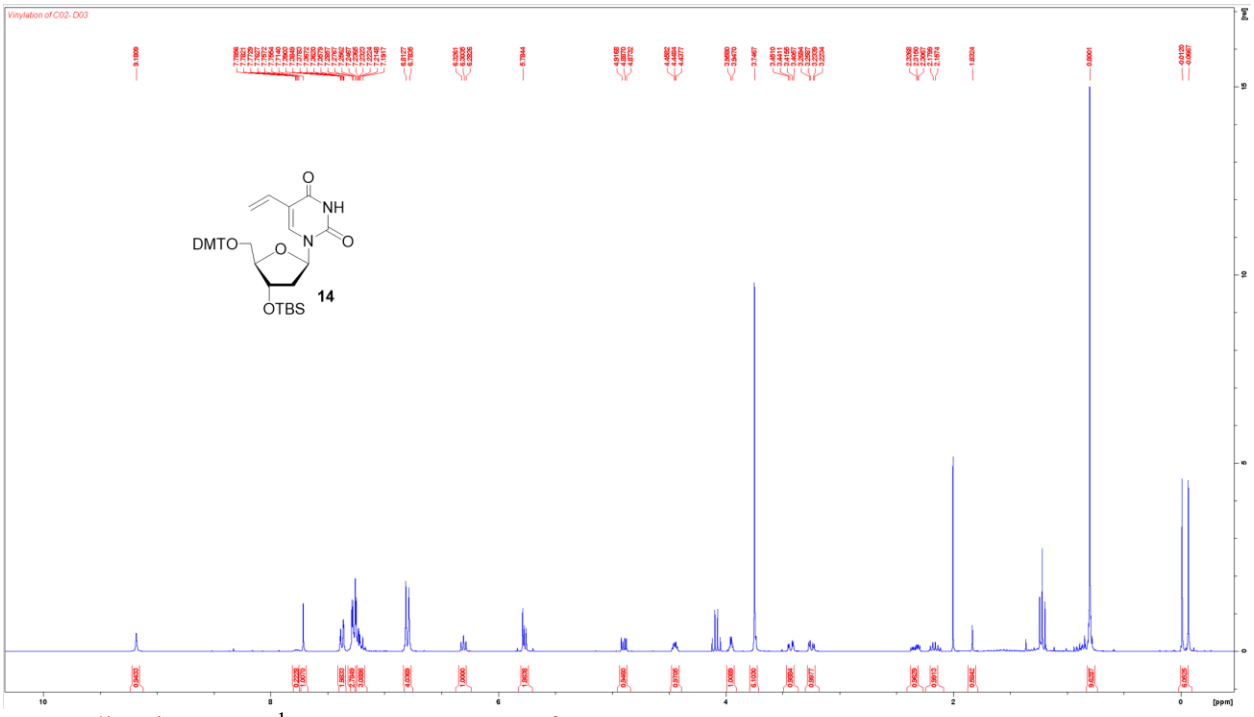
Appendix Figure 13. ^{1}H and ^{31}P NMR Spectrum of **12**.



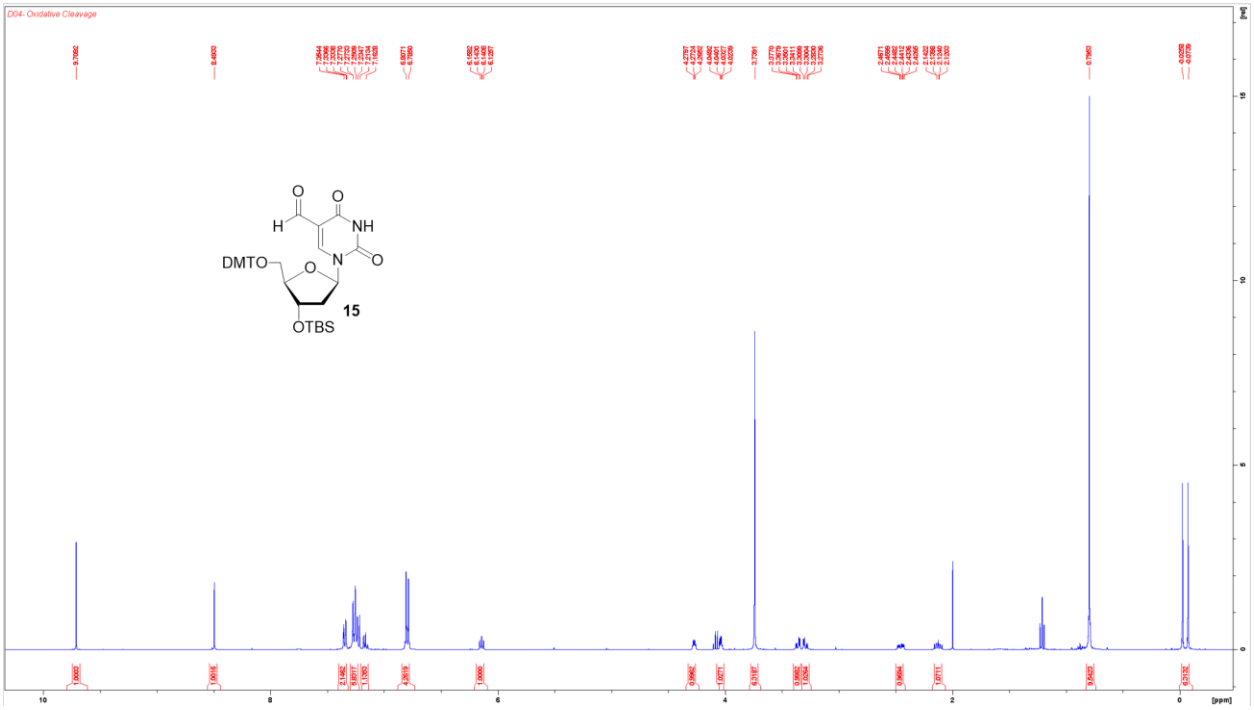
Appendix Figure 14. ^{1}H NMR Spectrum of Precursor to **13**.



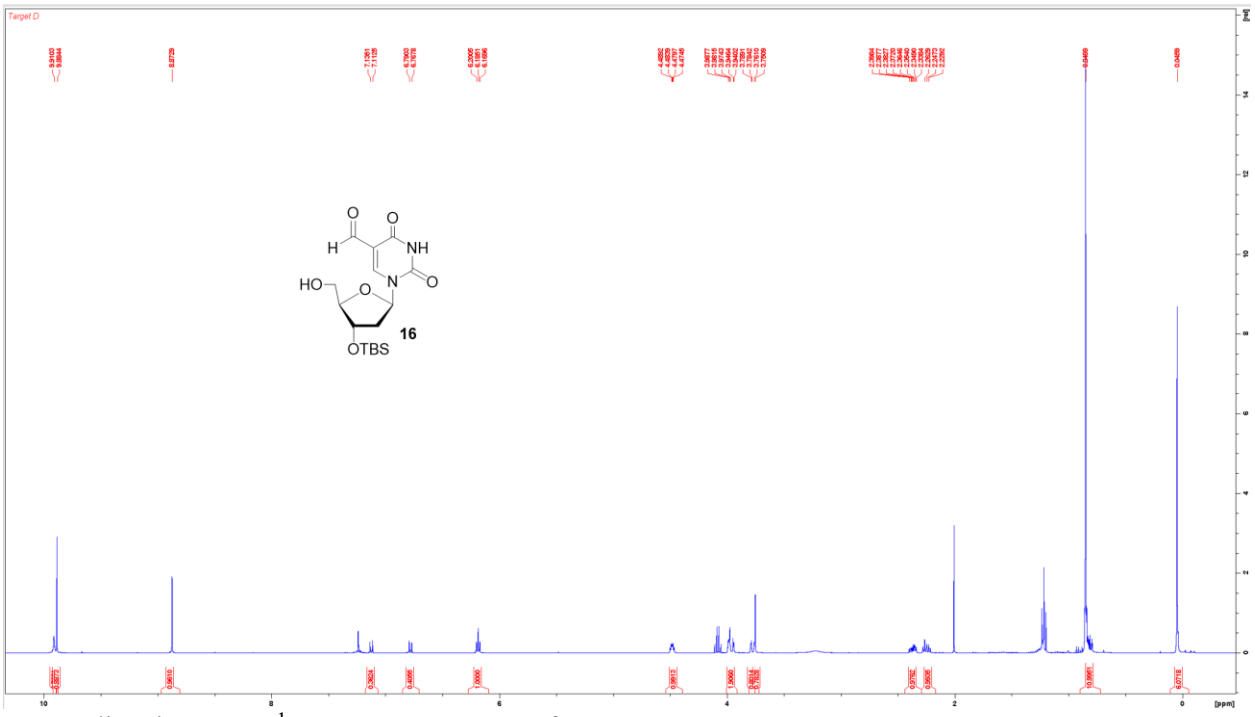
Appendix Figure 15. ^{1}H NMR Spectrum of **13**.



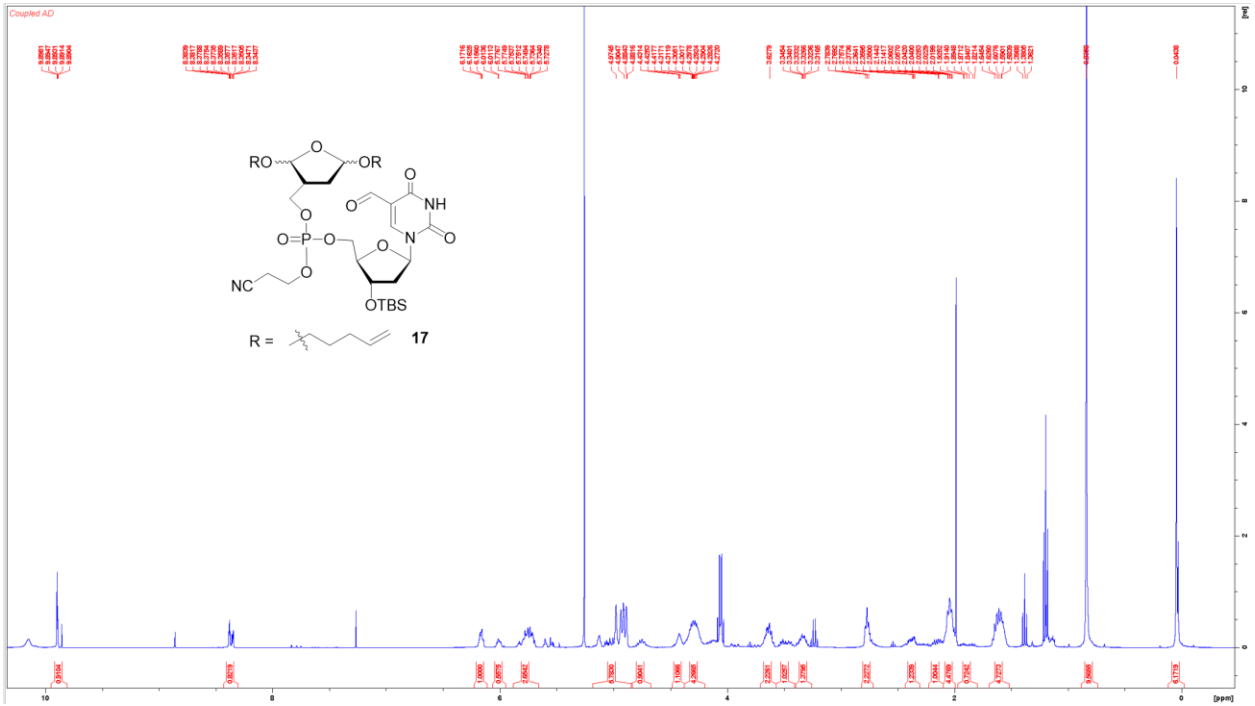
Appendix Figure 16. ^{1}H NMR Spectrum of **14**.



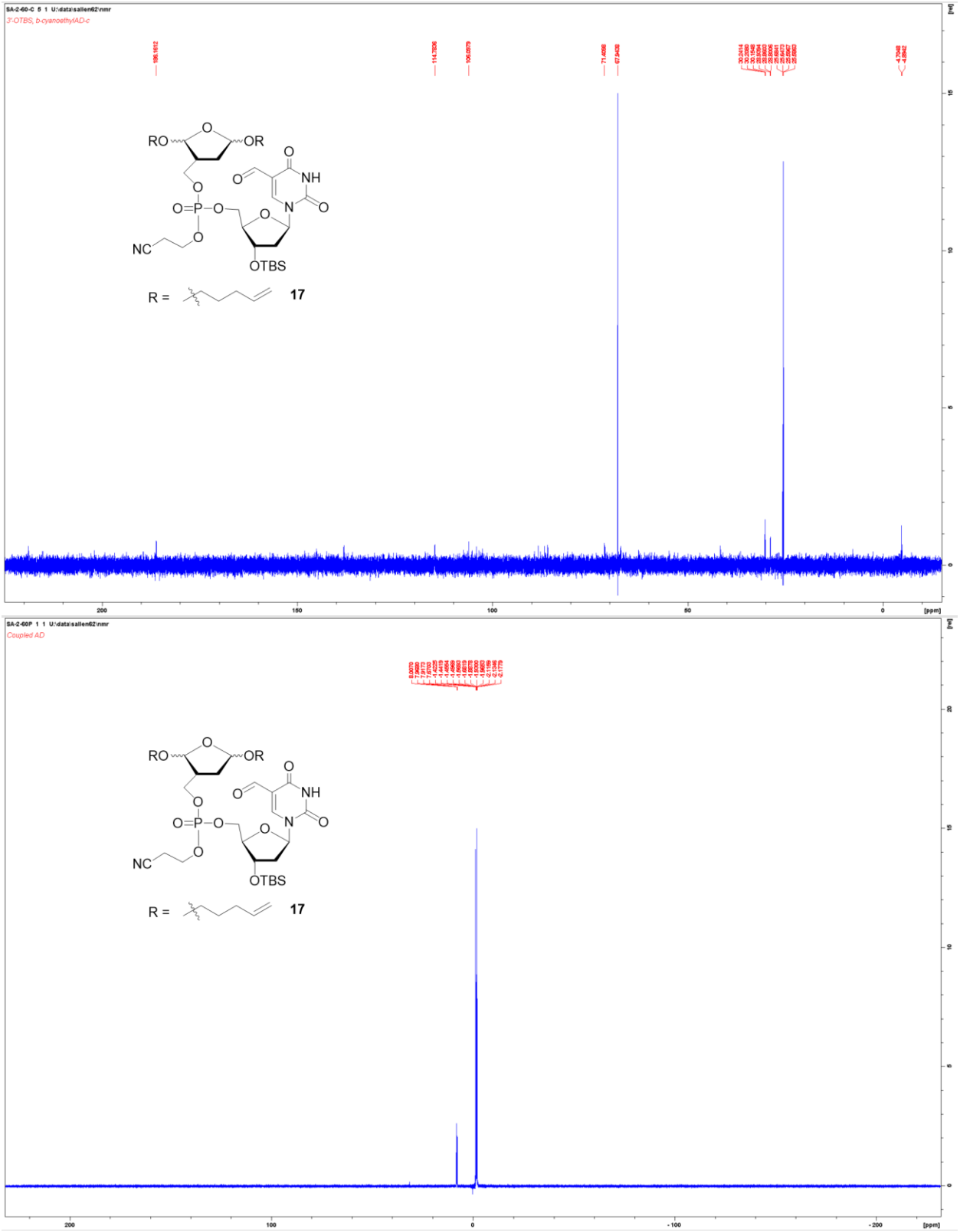
Appendix Figure 17. ^{1}H NMR Spectrum of **15**.



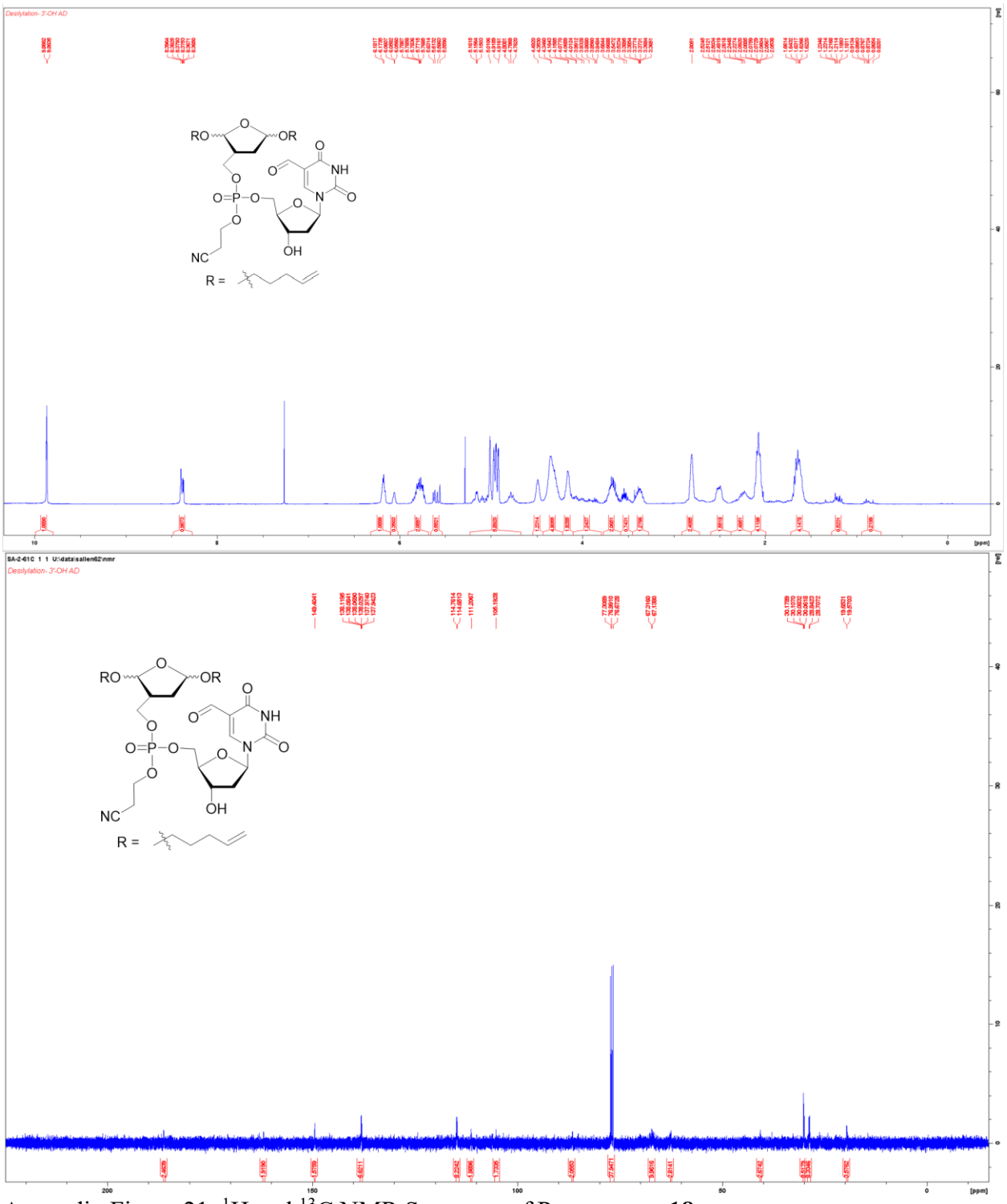
Appendix Figure 18. ^{1}H NMR Spectrum of **16**.



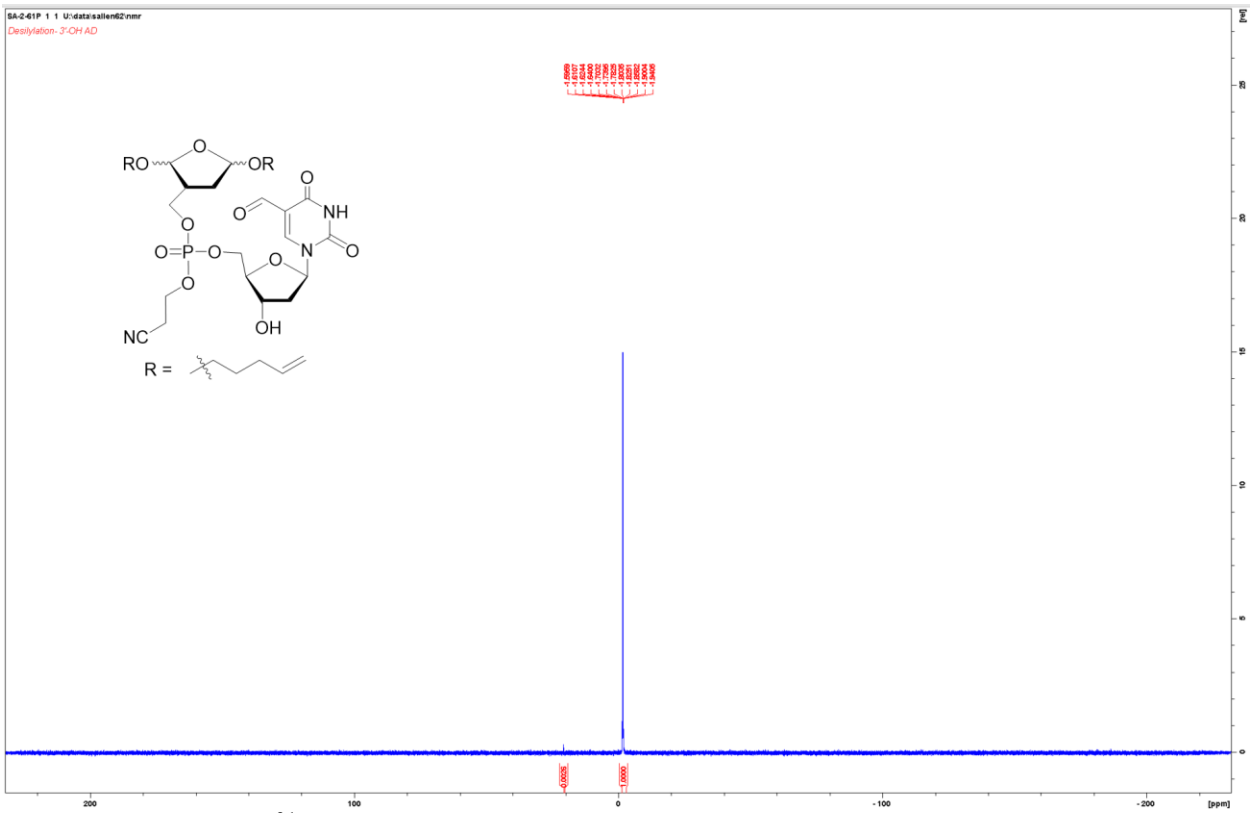
Appendix Figure 19. ^{1}H NMR Spectrum of **17**.



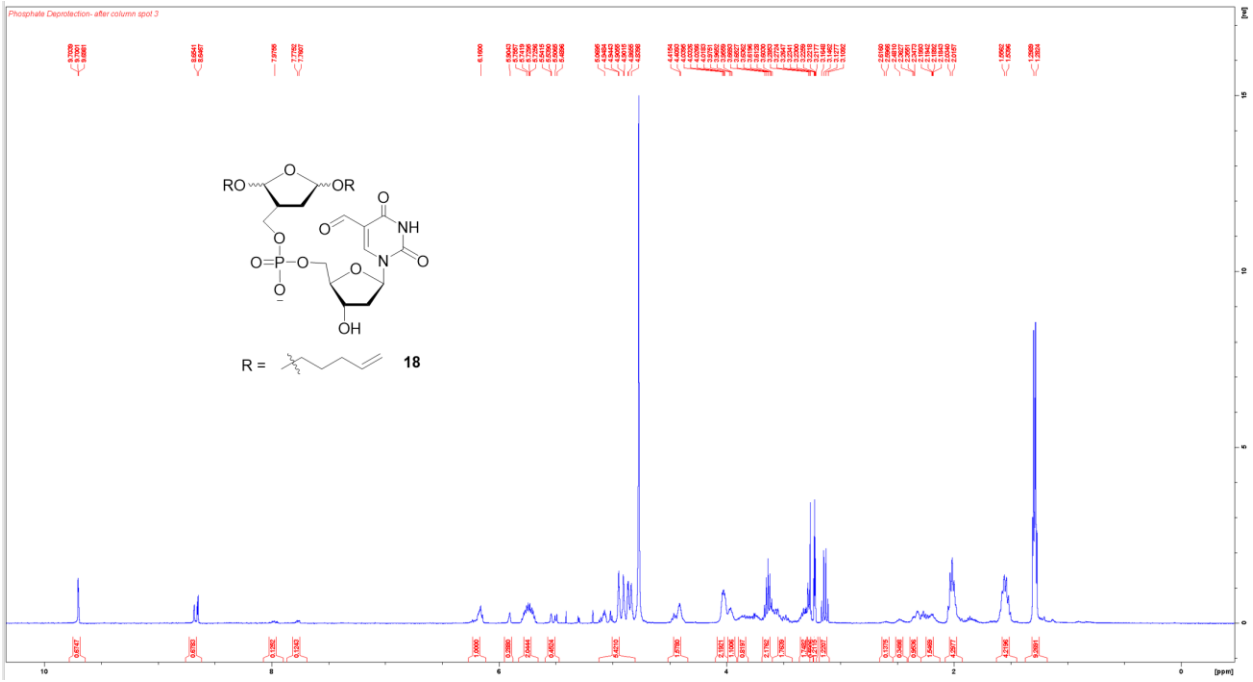
Appendix Figure 20. ^{13}C and ^{31}P NMR Spectrum of **17**.



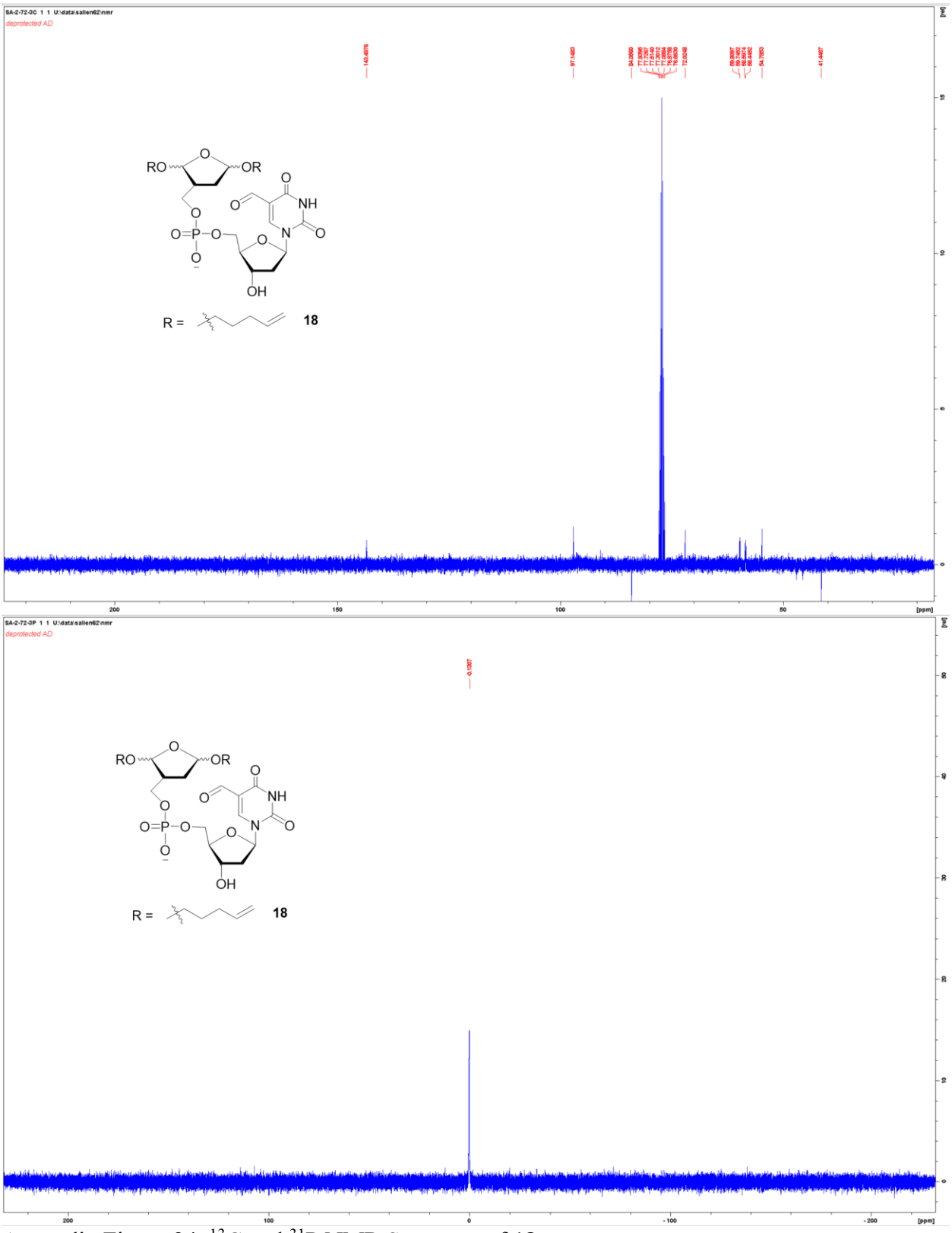
Appendix Figure 21. ^{1}H and ^{13}C NMR Spectrum of Precursor to **18**.



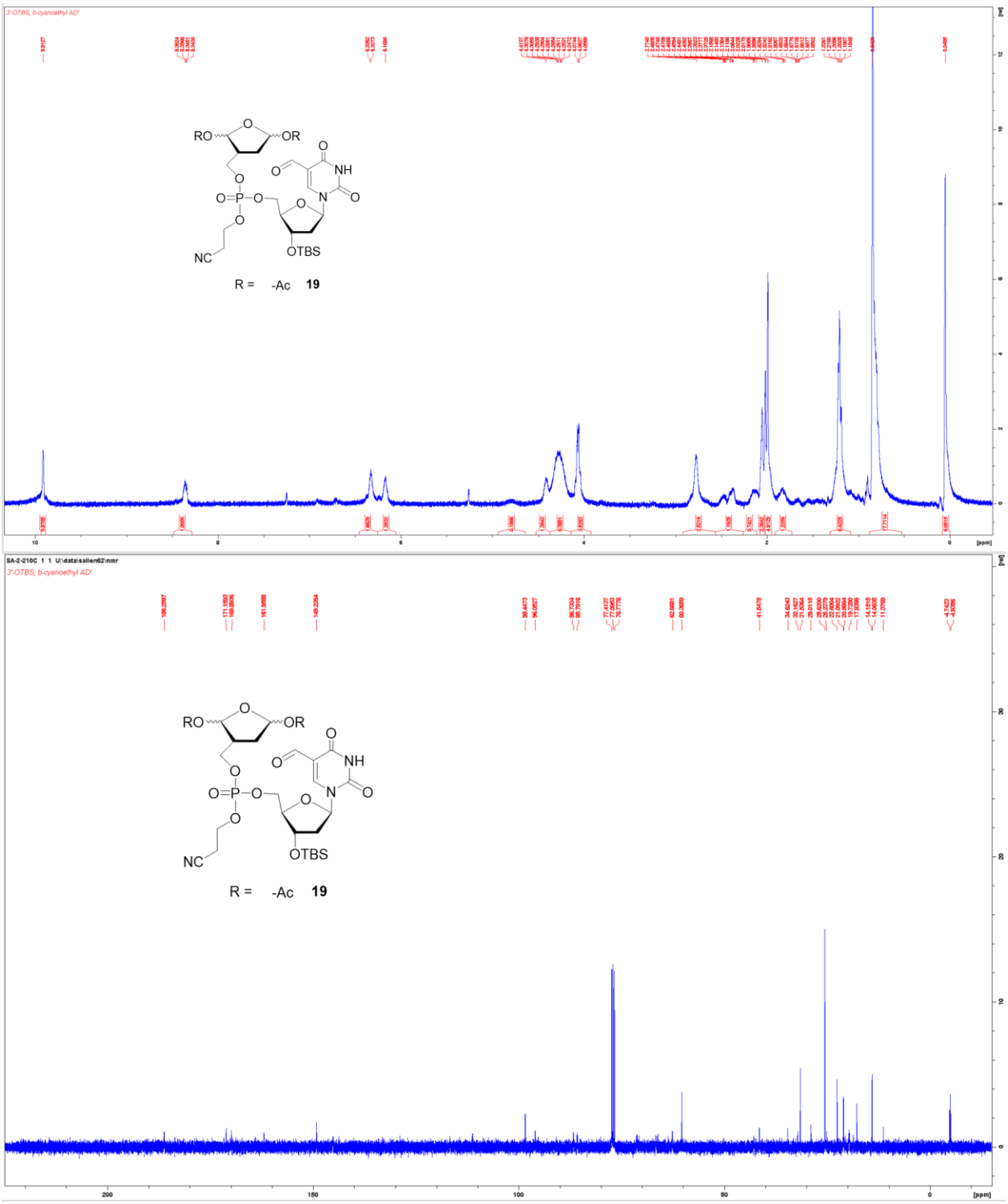
Appendix Figure 22. ^{31}P NMR Spectrum of Precursor to **18**.



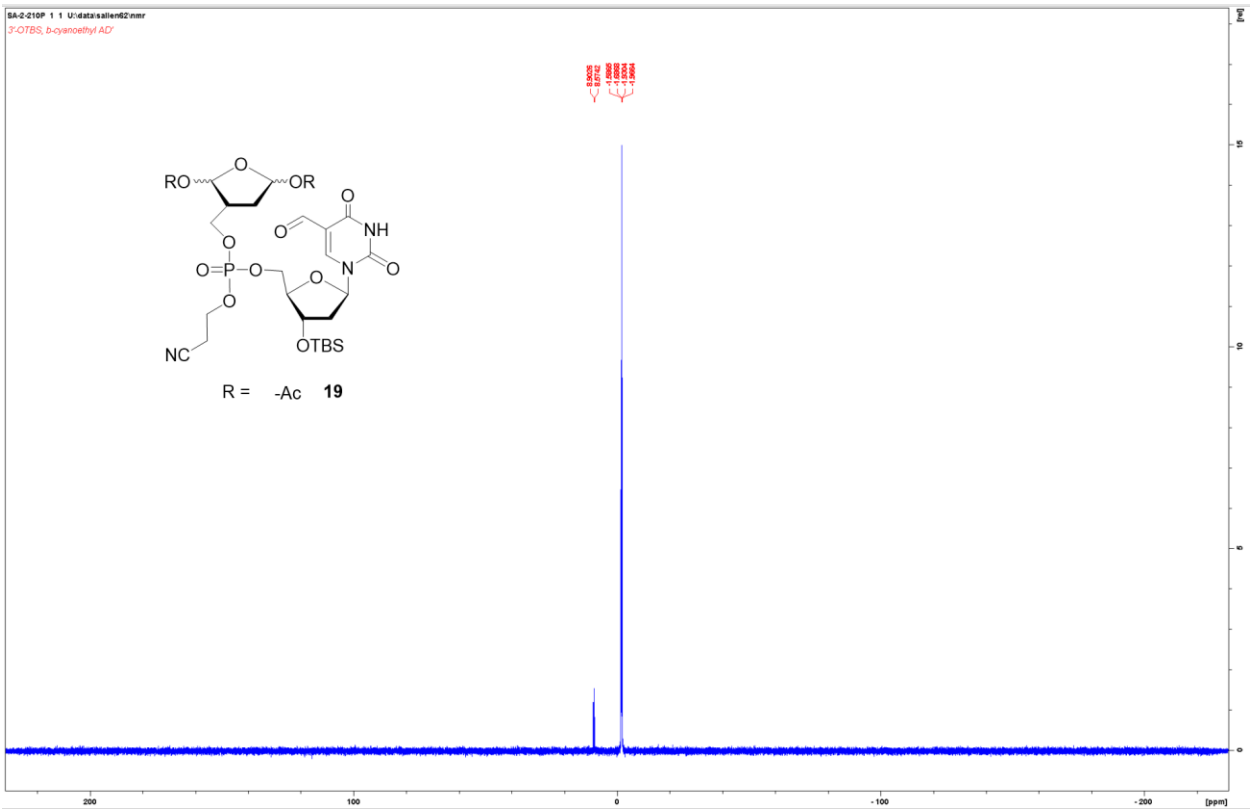
Appendix Figure 23. ^{1}H NMR Spectrum of **18**.



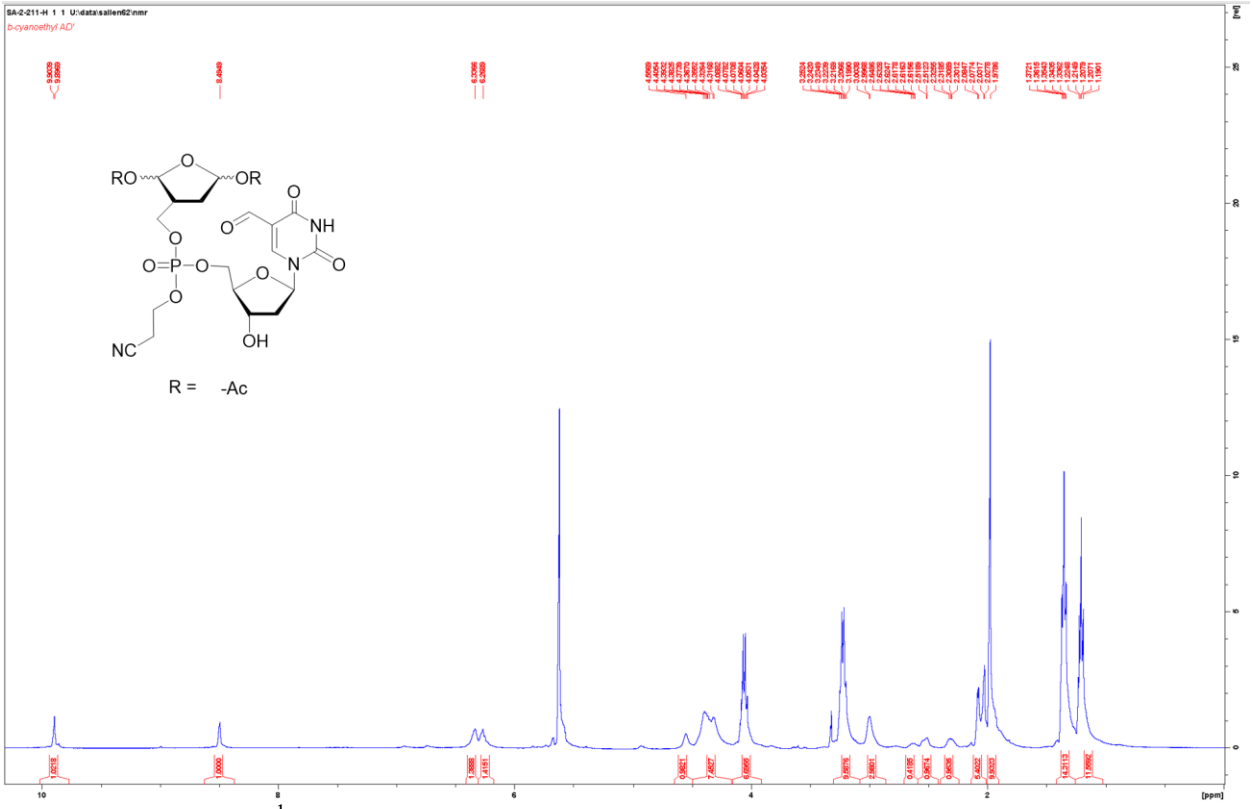
Appendix Figure 24. ^{13}C and ^{31}P NMR Spectrum of **18**.



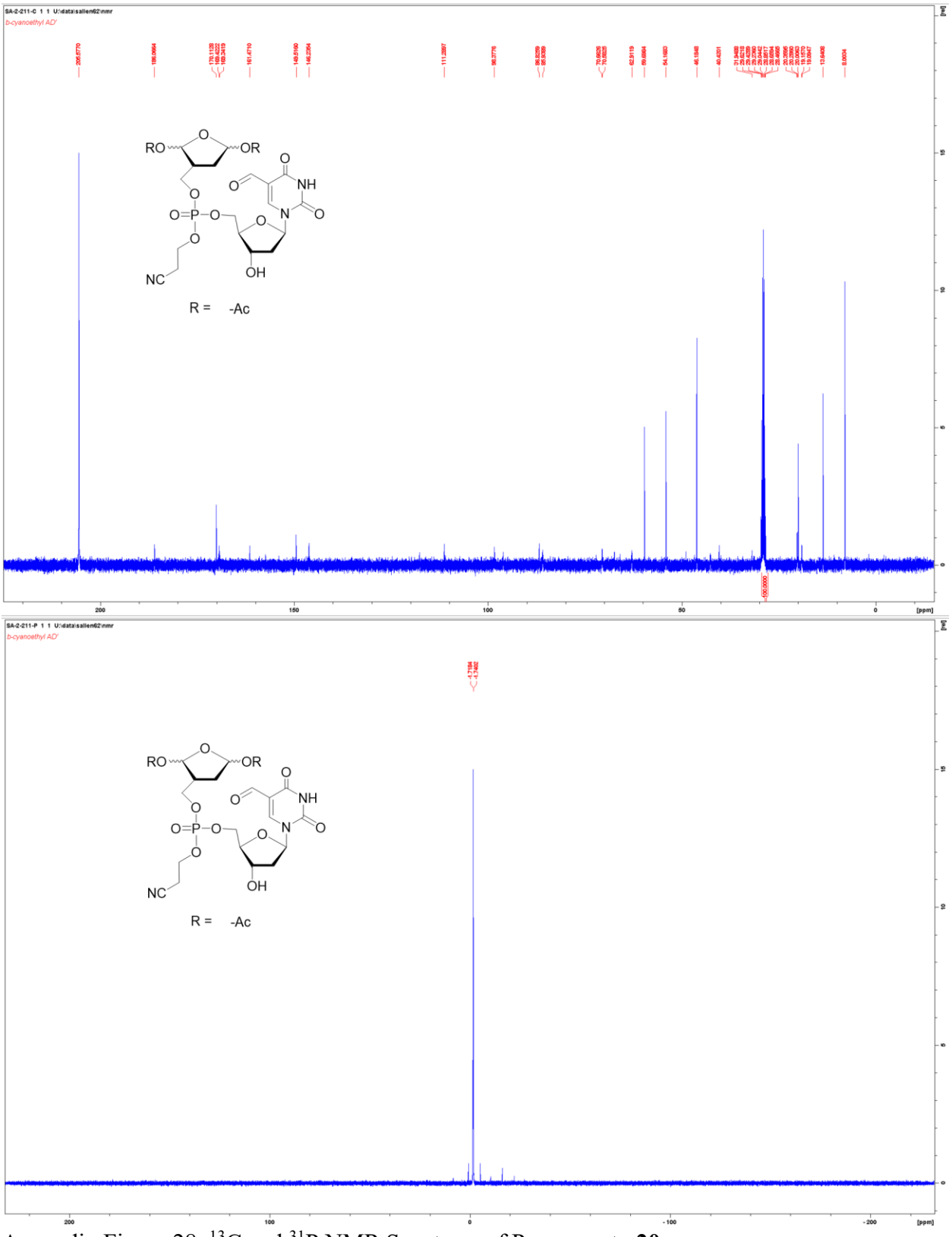
Appendix Figure 25. ^{1}H and ^{13}C NMR Spectrum of **19**.



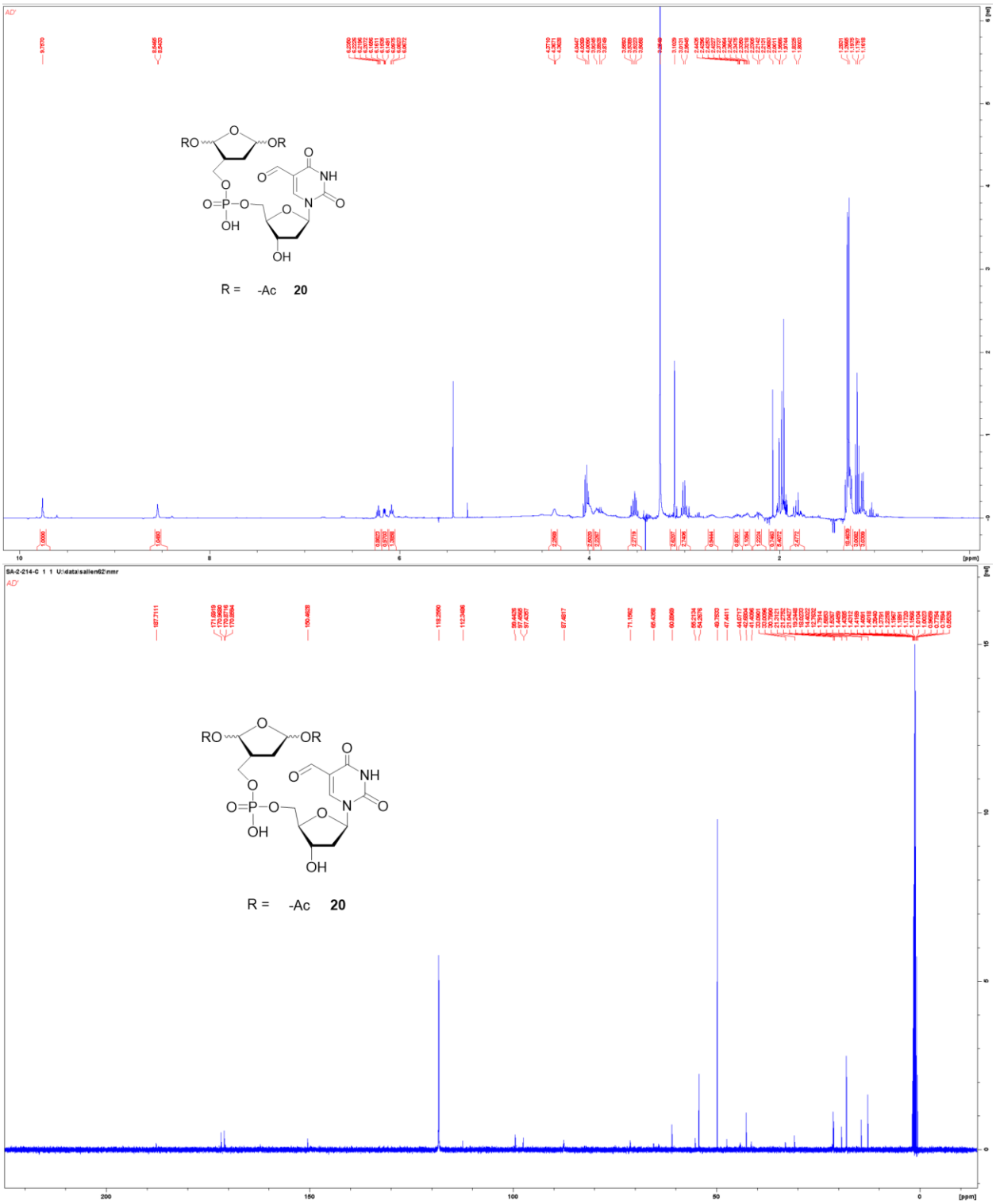
Appendix Figure 26. ^{31}P NMR Spectrum of **19**.



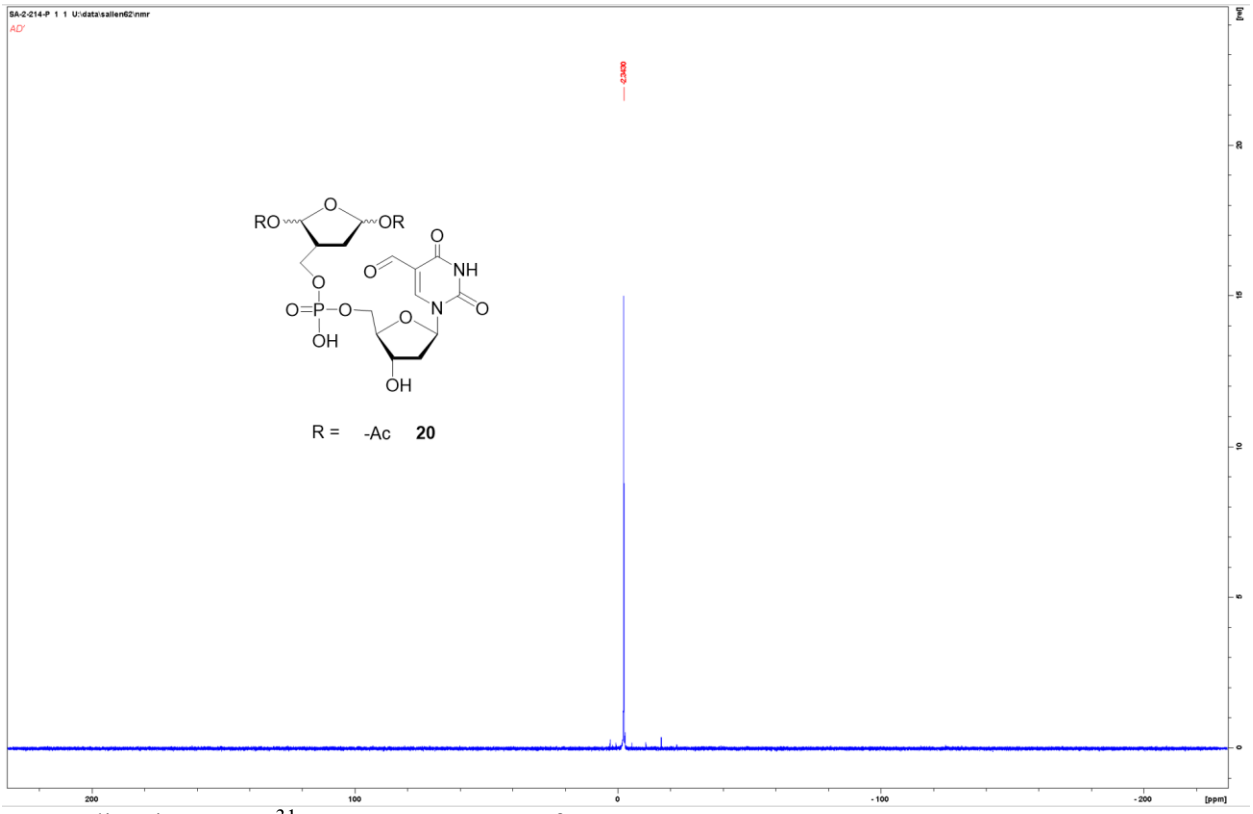
Appendix Figure 27. ^{1}H NMR Spectrum of Precursor to **20**.



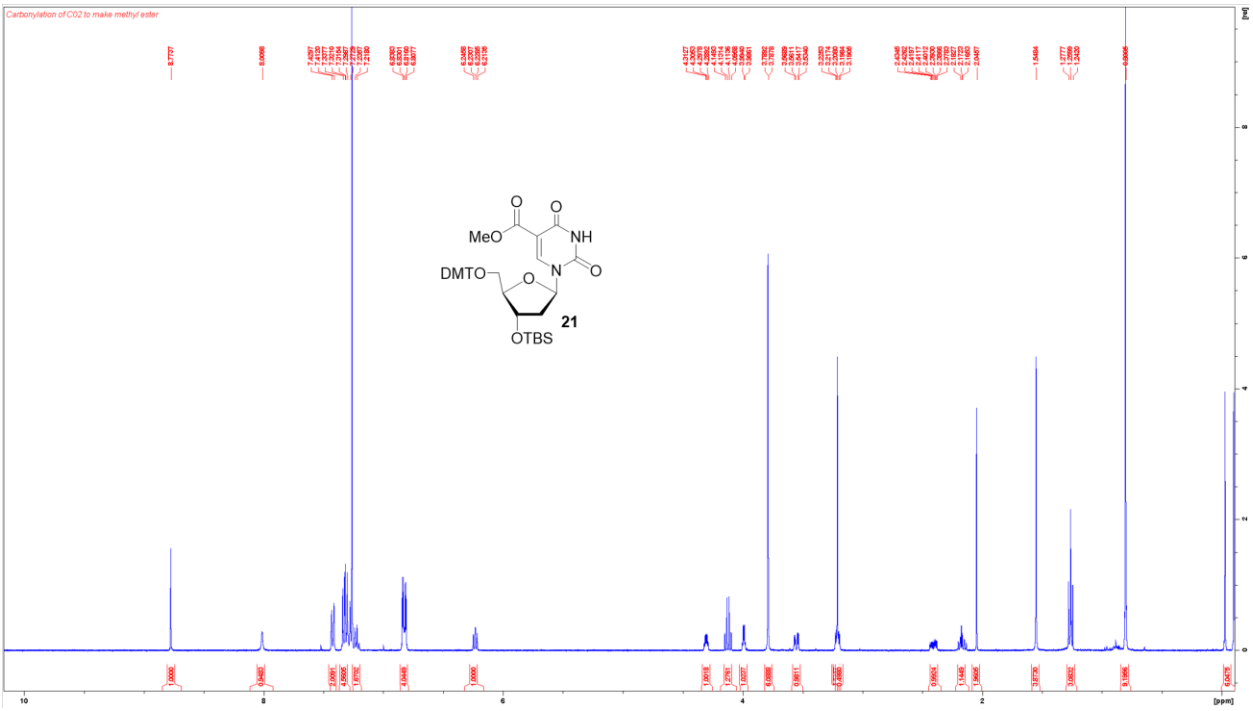
Appendix Figure 28. ^{13}C and ^{31}P NMR Spectrum of Precursor to **20**.



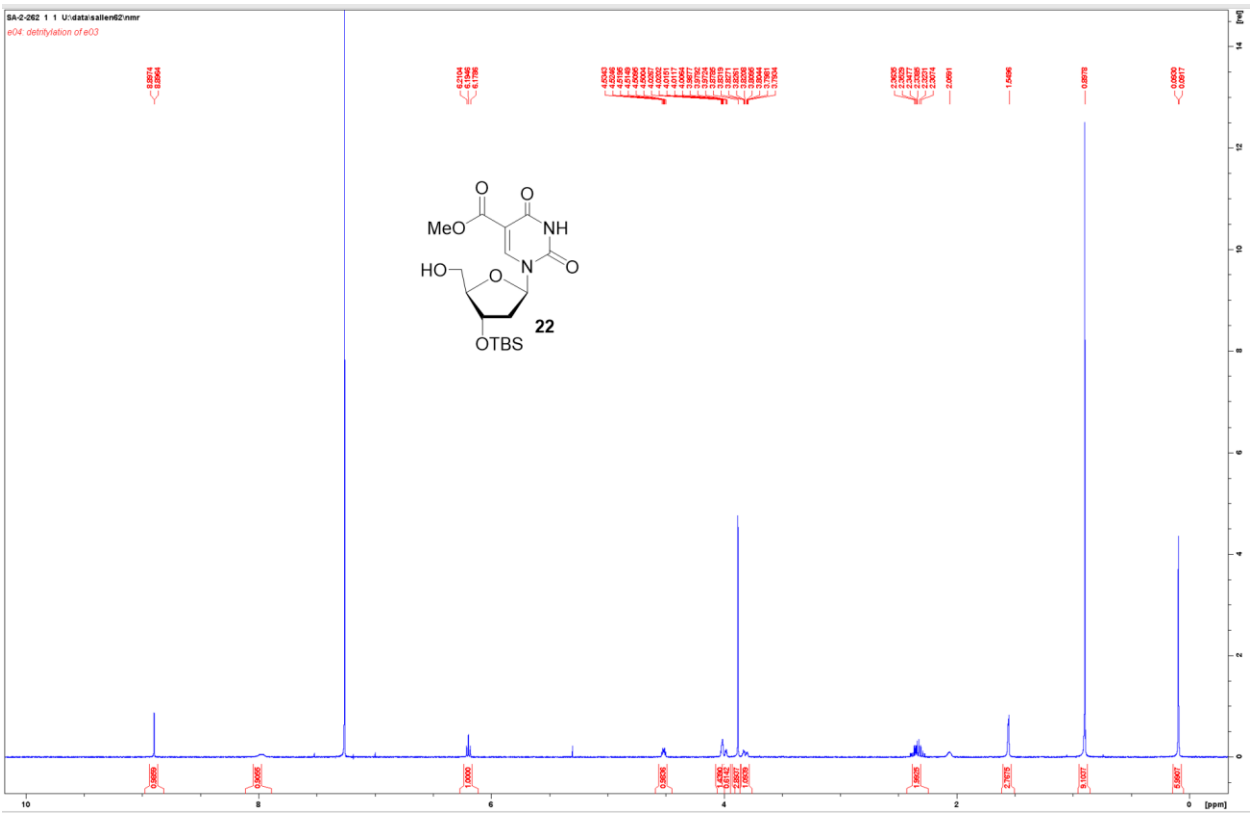
Appendix Figure 29. ^{1}H and ^{13}C NMR Spectrum of **20**.



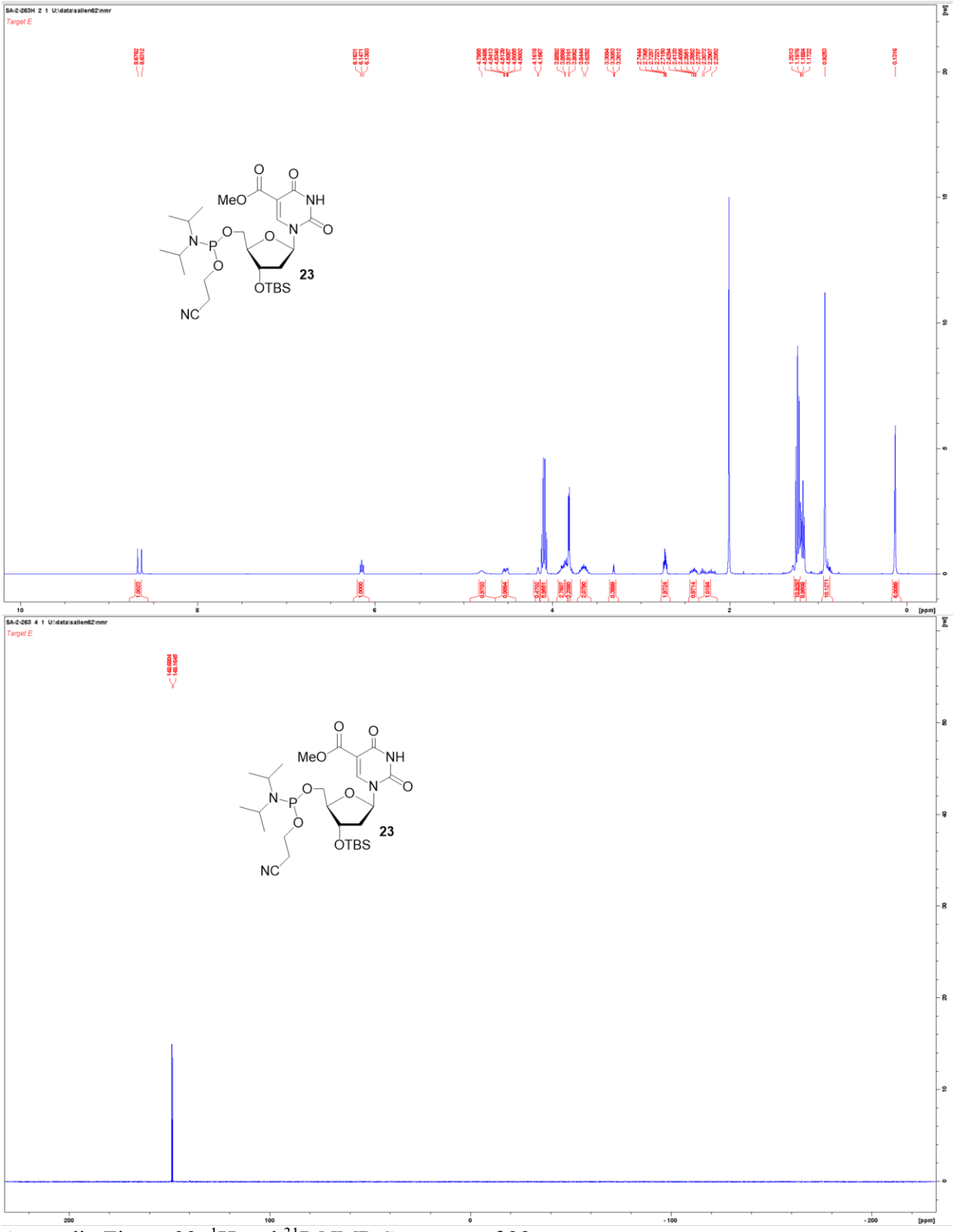
Appendix Figure 30. ^{31}P NMR Spectrum of **20**.



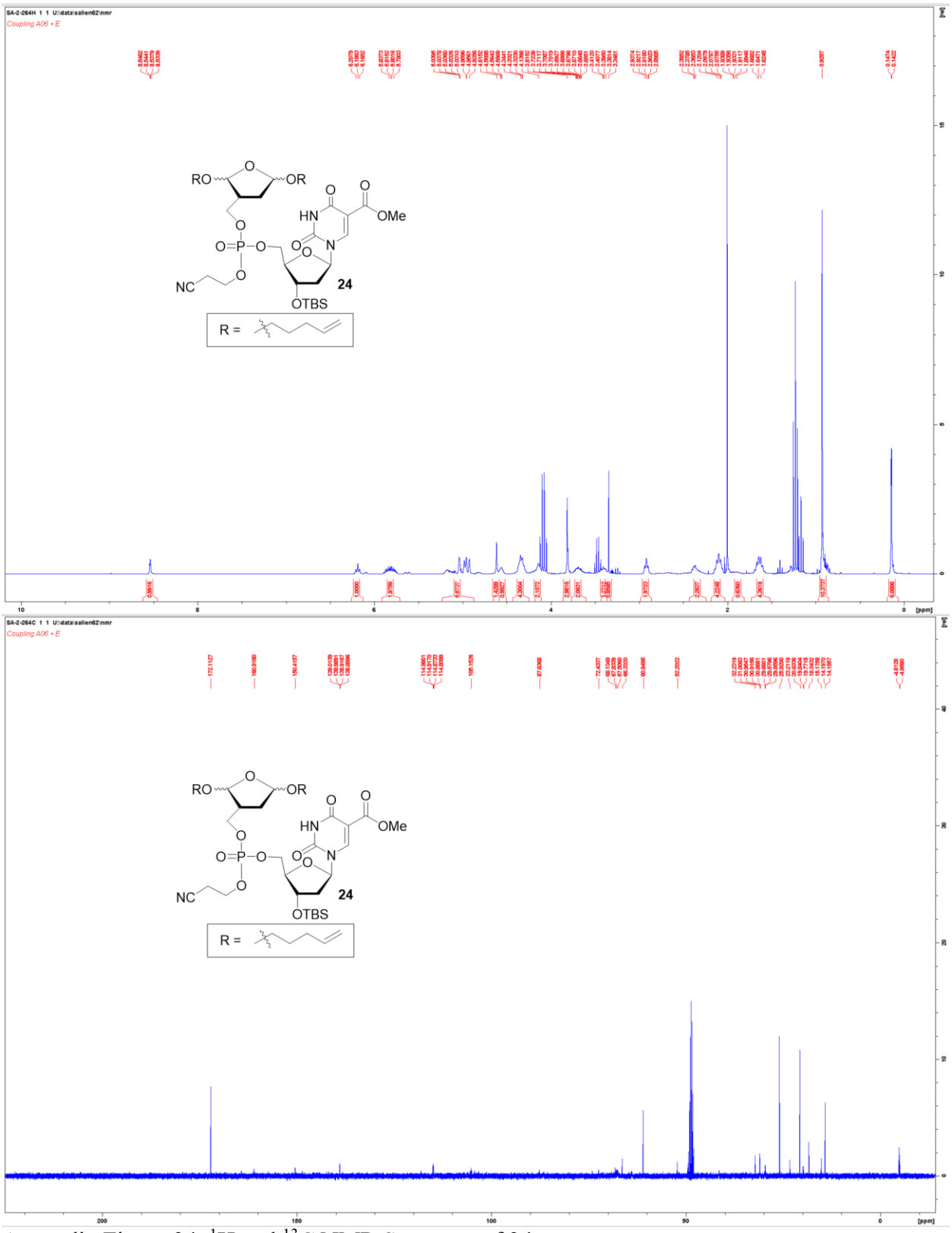
Appendix Figure 31. ^{1}H NMR Spectrum of **21**.



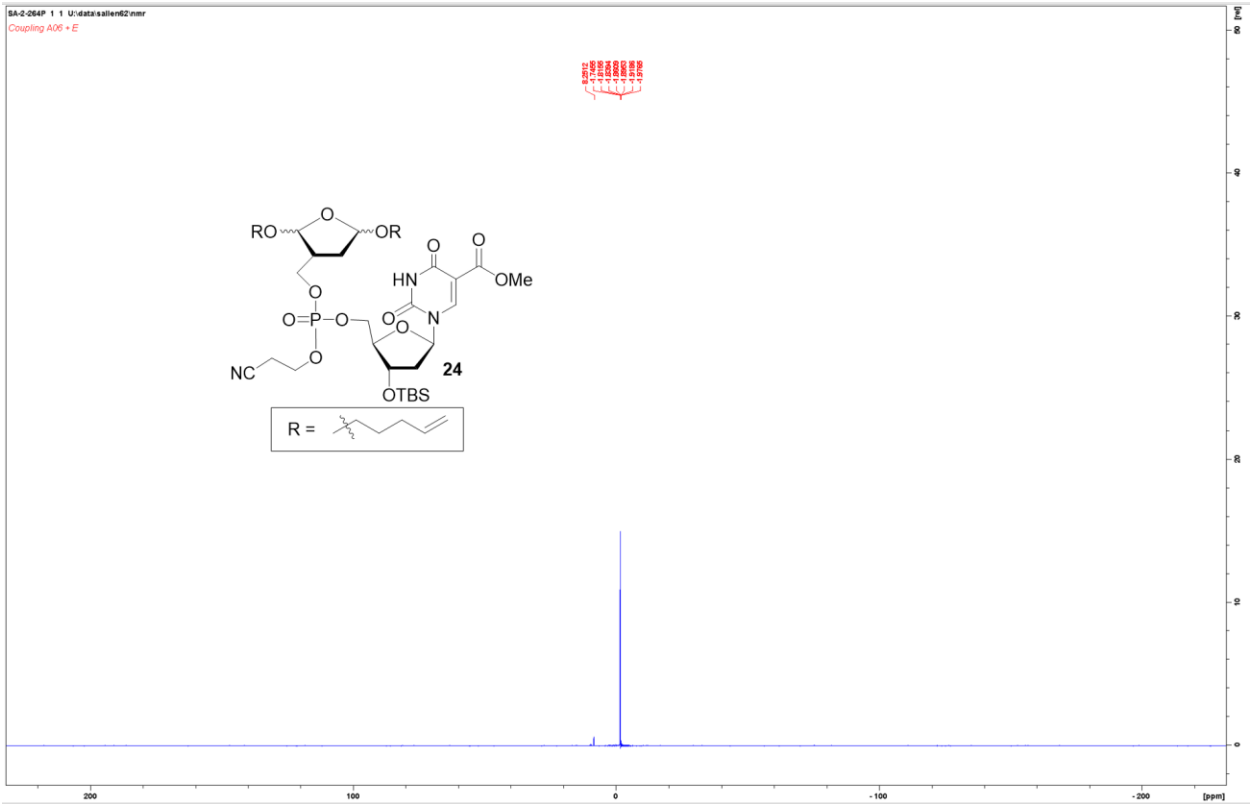
Appendix Figure 32. ^{1}H NMR Spectrum of **22**.



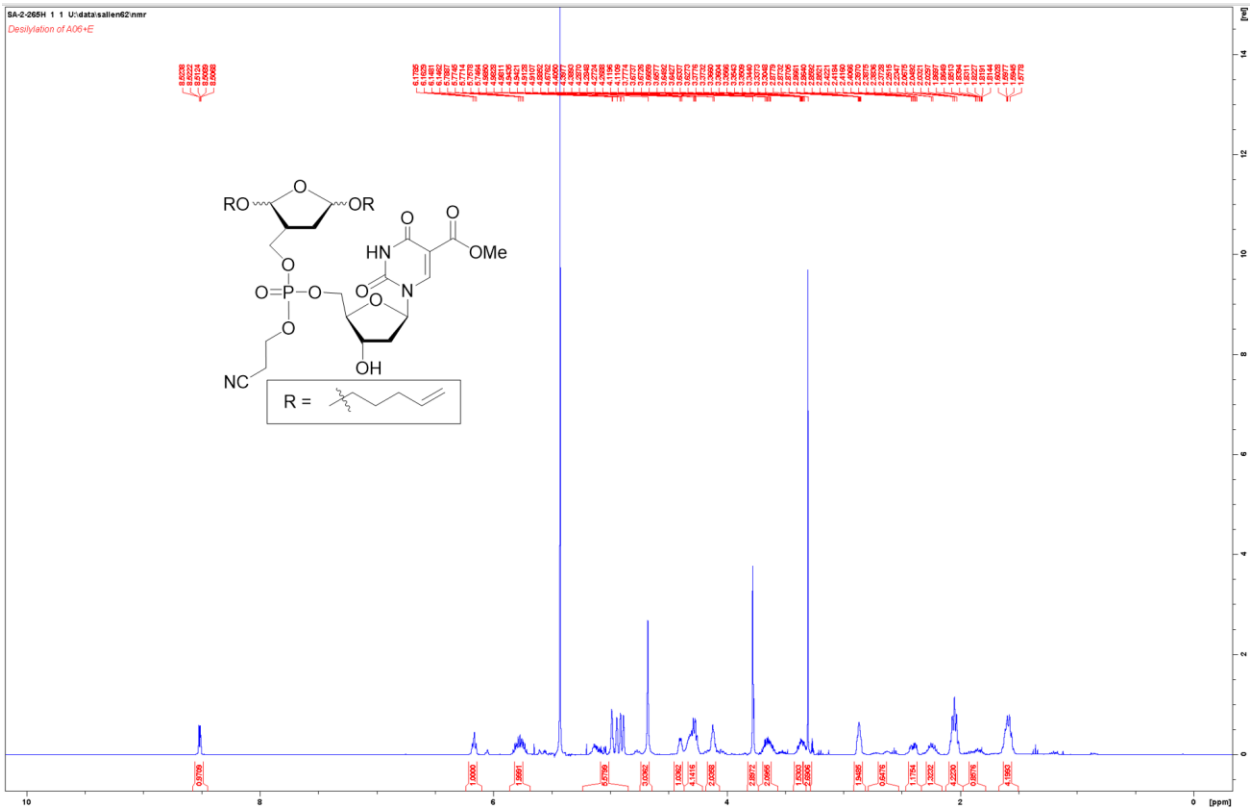
Appendix Figure 33. ^{1}H and ^{31}P NMR Spectrum of **23**.



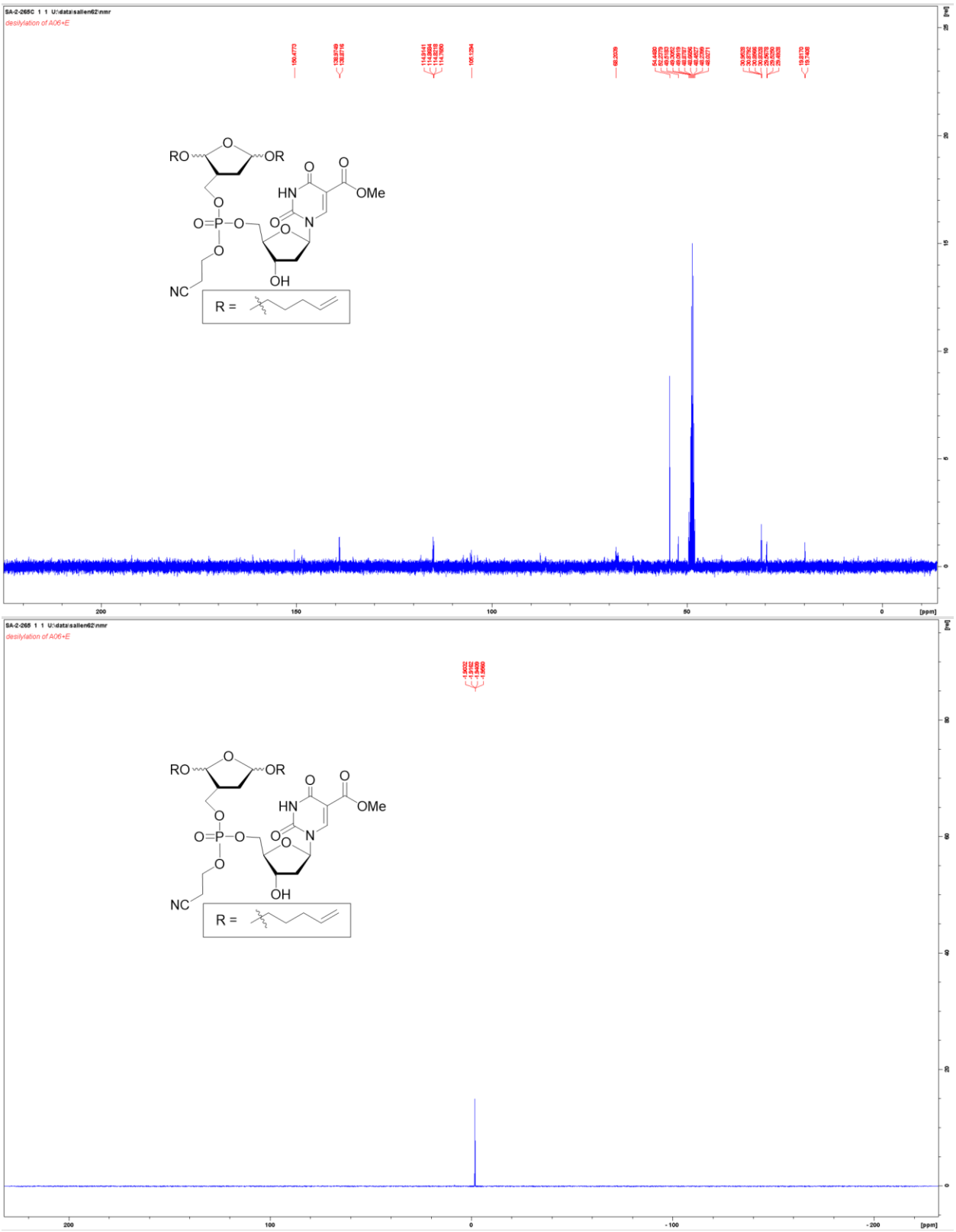
Appendix Figure 34. ^{1}H and ^{13}C NMR Spectrum of **24**.



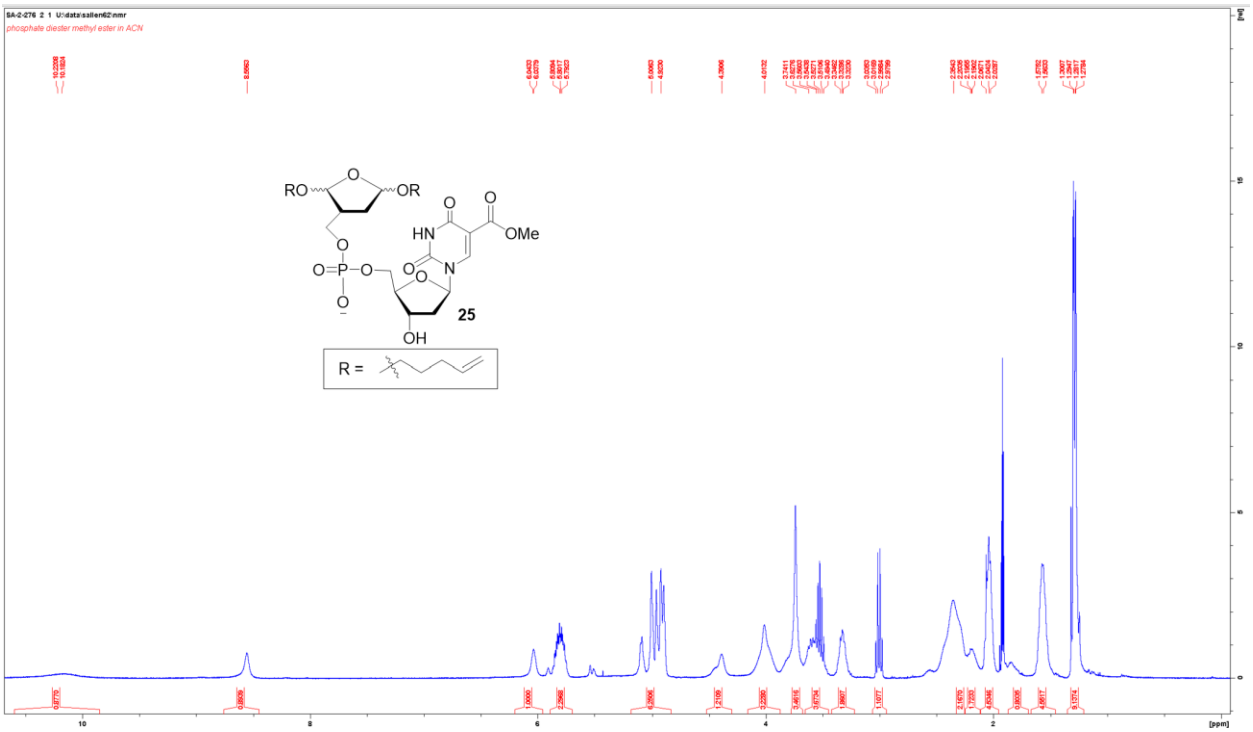
Appendix Figure 35. ^{31}P NMR Spectrum of **24**.



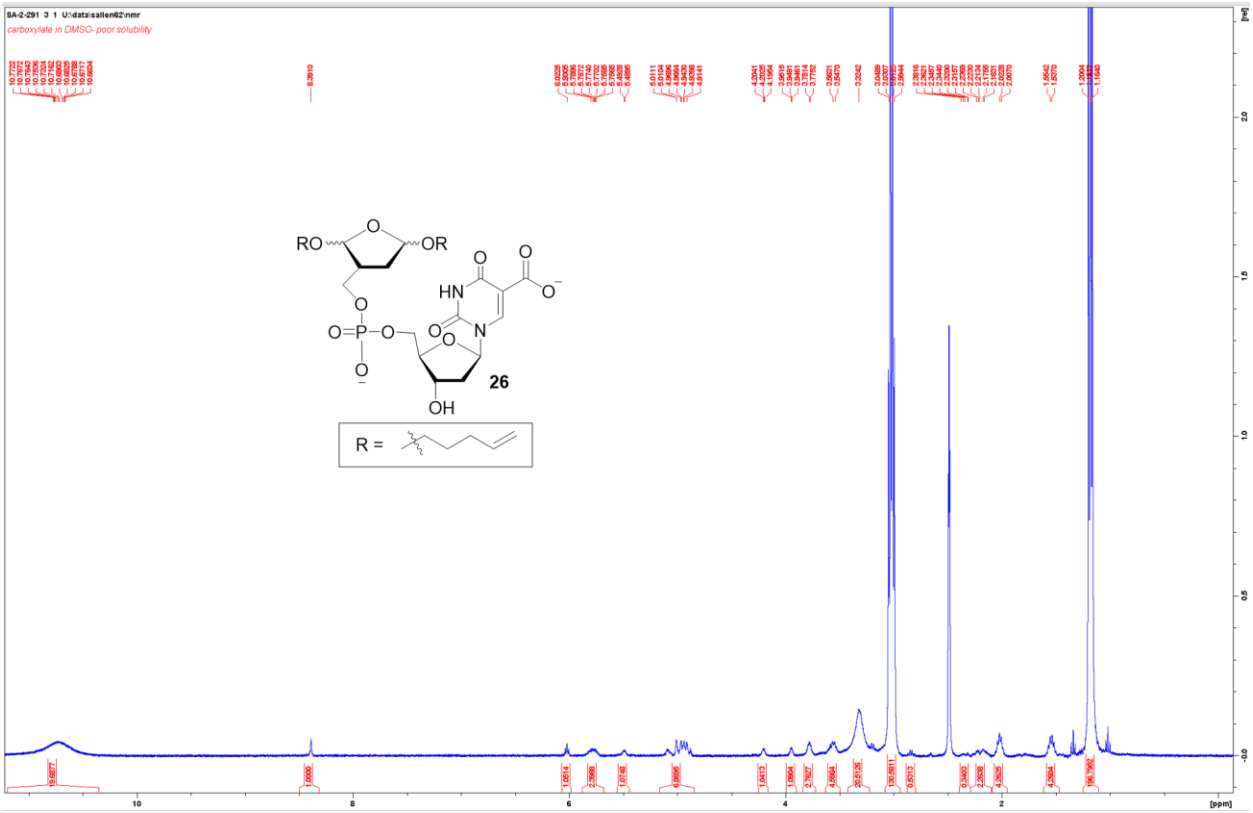
Appendix Figure 36. ^{1}H NMR Spectrum of Precursor to **25**.



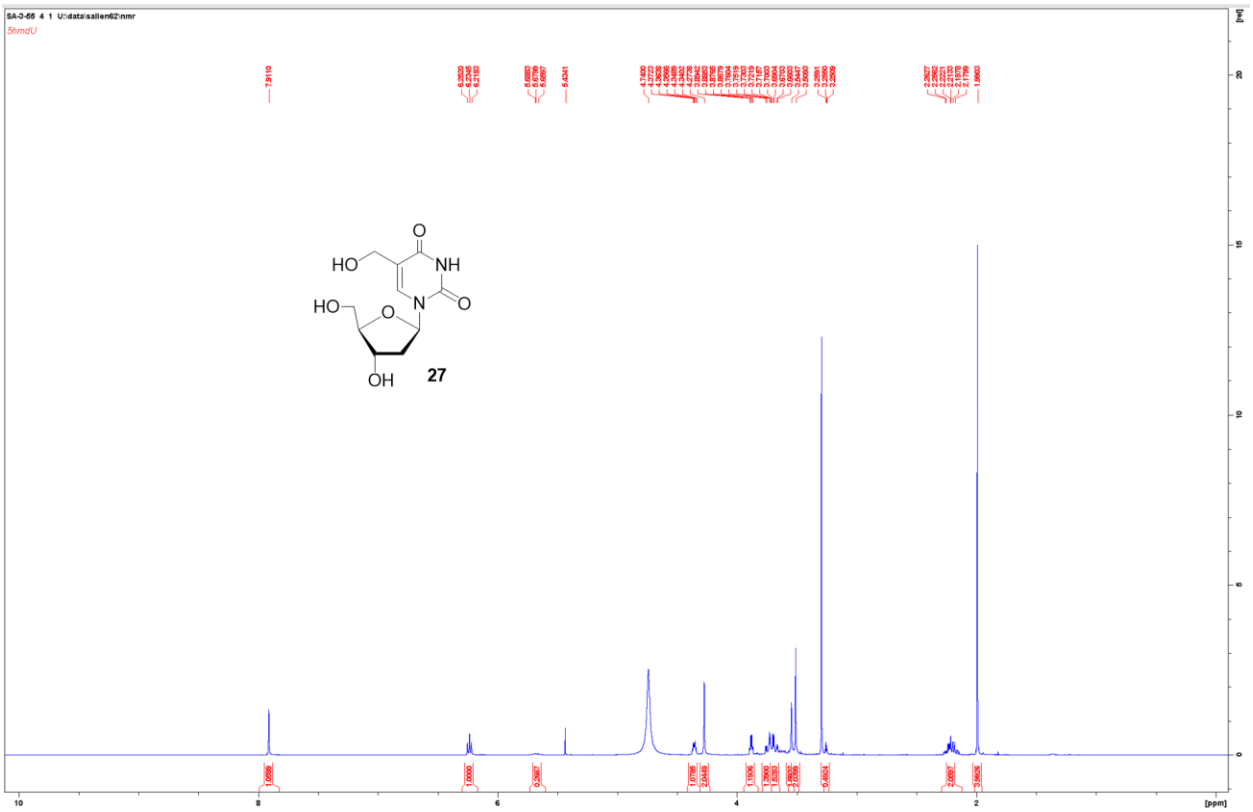
Appendix Figure 37. ^{13}C and ^{31}P NMR Spectrum of Precursor to **25**.



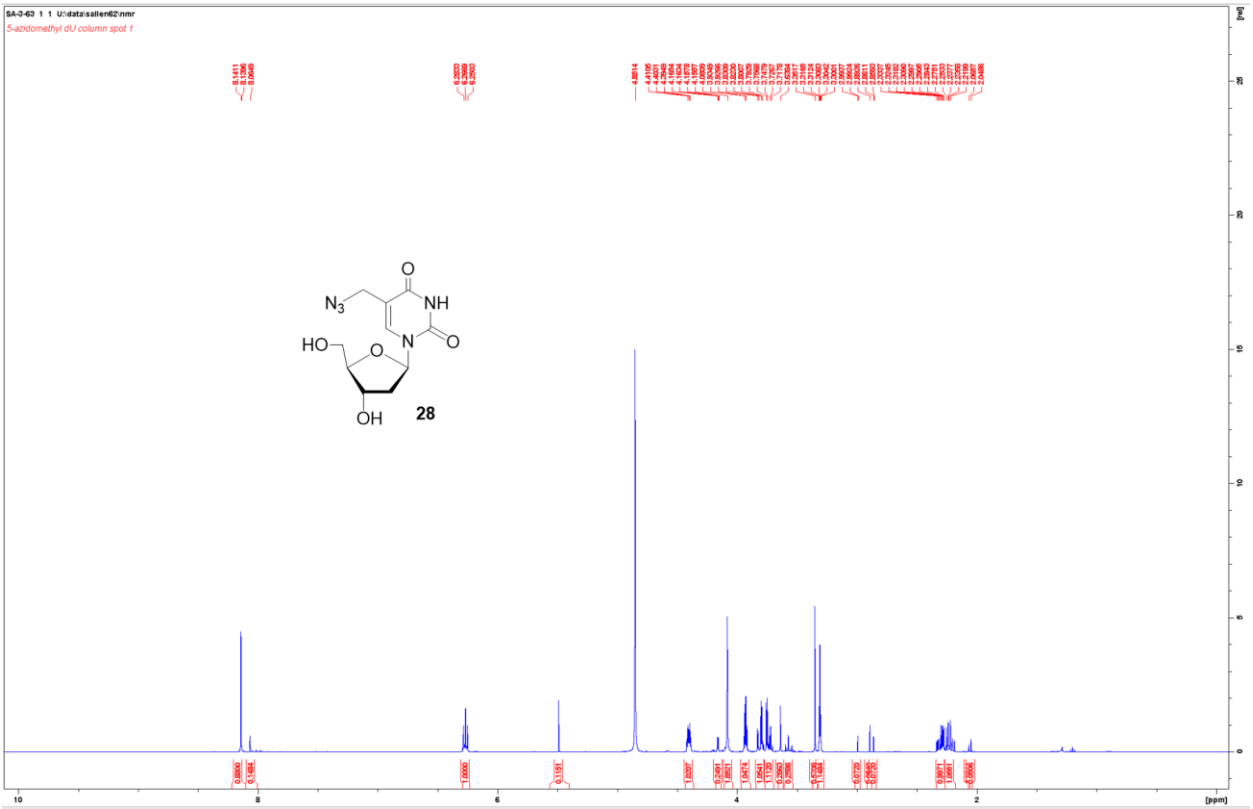
Appendix Figure 38. ^{1}H NMR Spectrum of **25**.



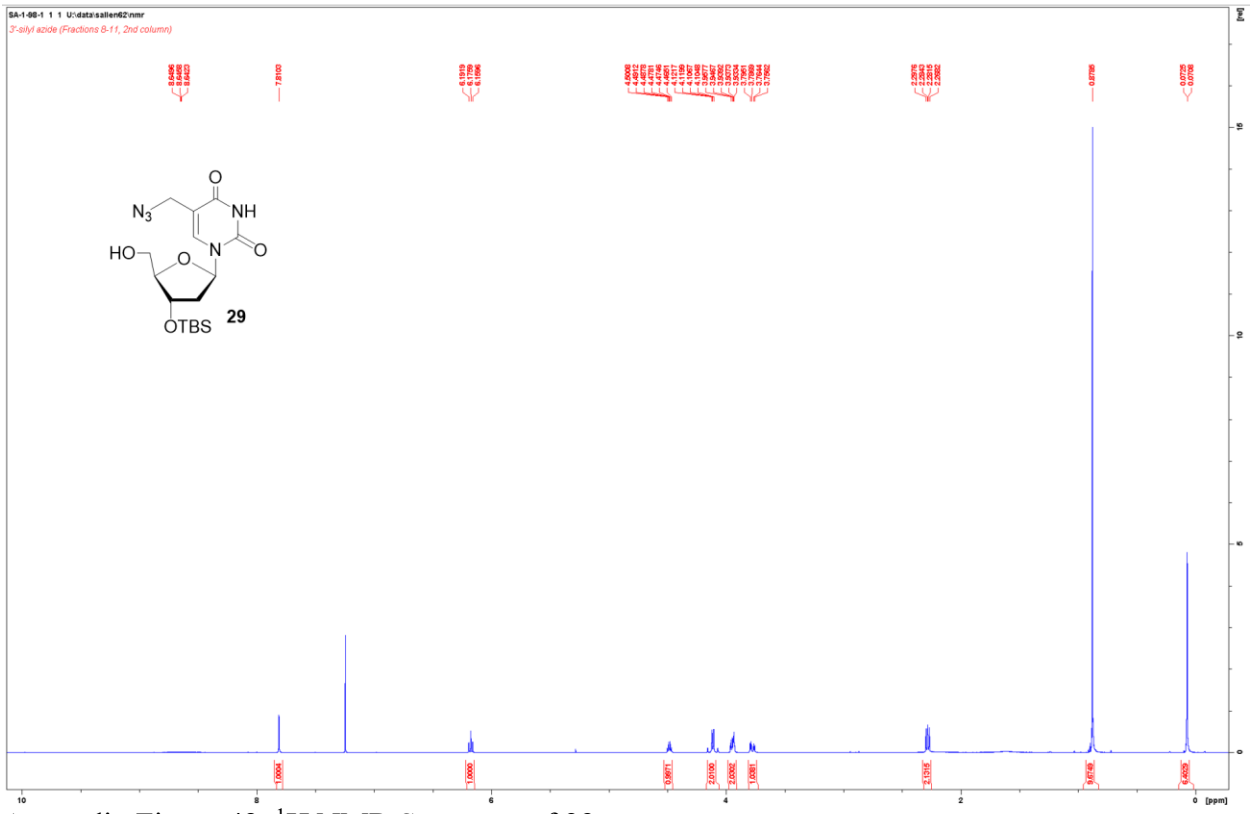
Appendix Figure 39. ^{1}H NMR Spectrum of **26**.



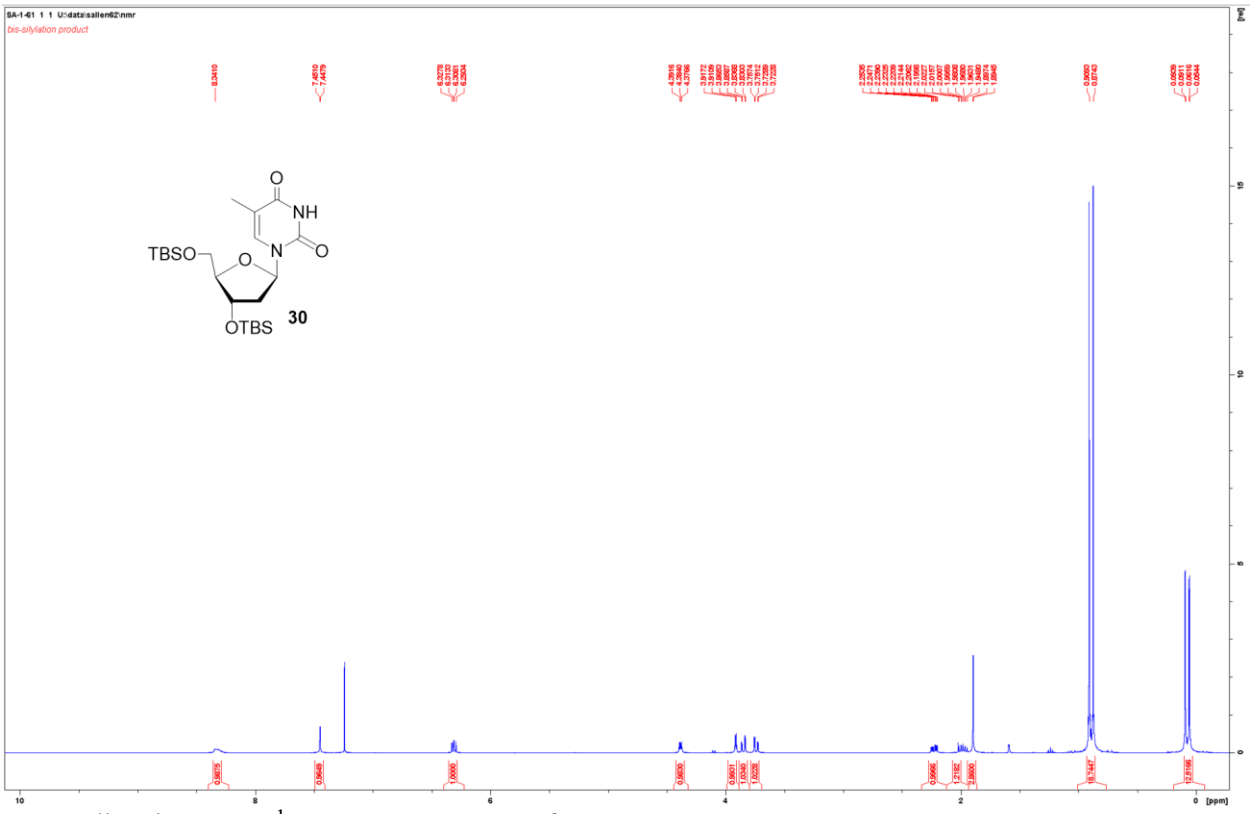
Appendix Figure 40. ^{1}H NMR Spectrum of **27**.



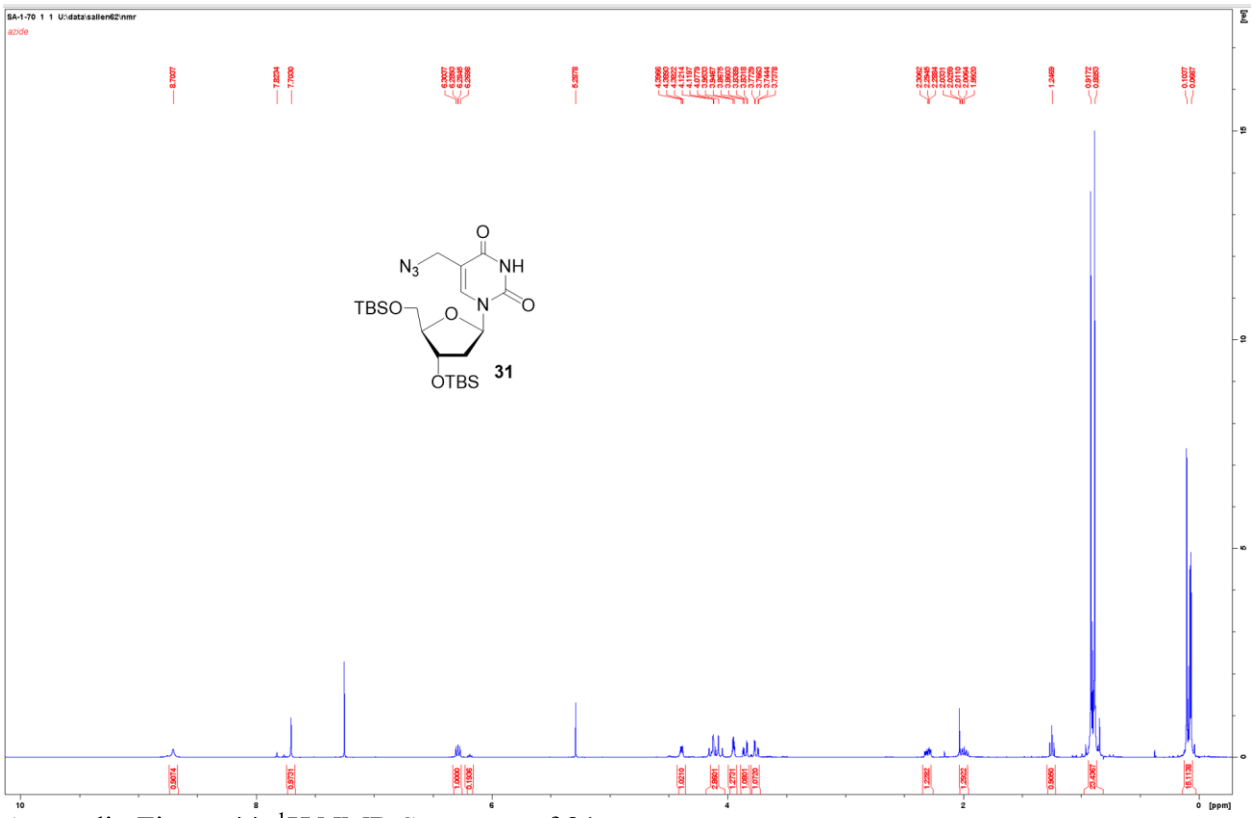
Appendix Figure 41. ^{1}H NMR Spectrum of **28**.



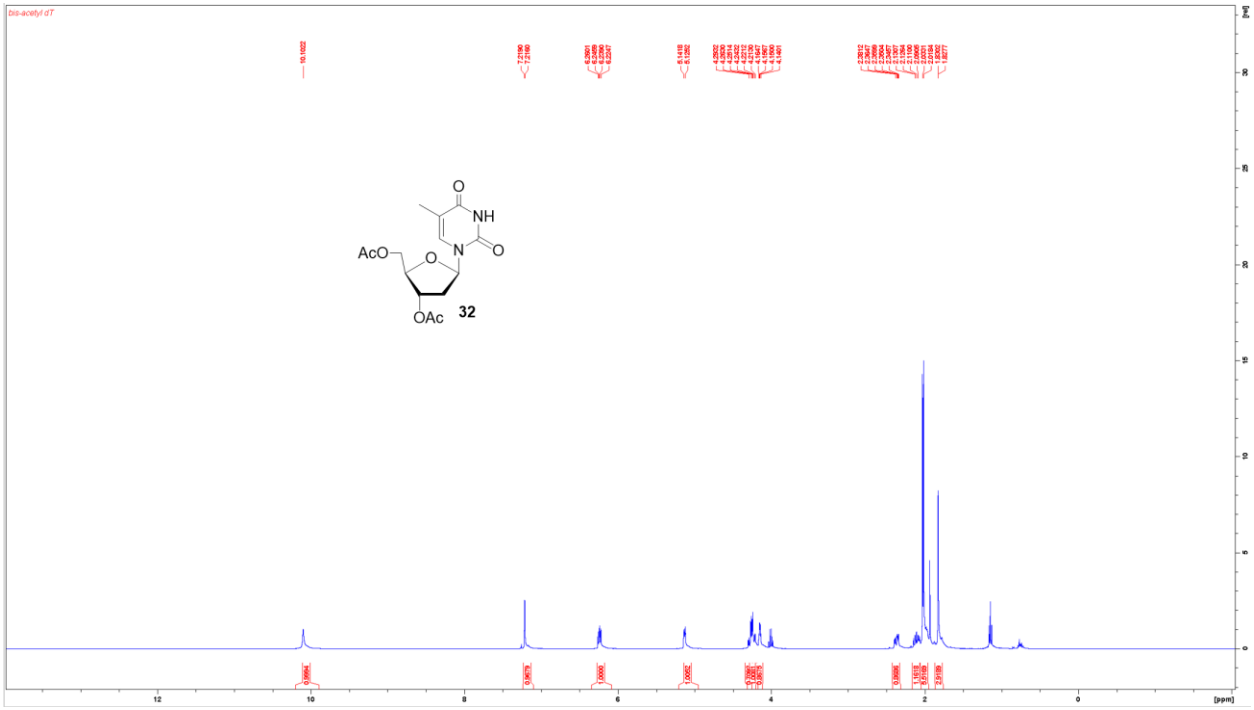
Appendix Figure 42. ^{1}H NMR Spectrum of **29**.



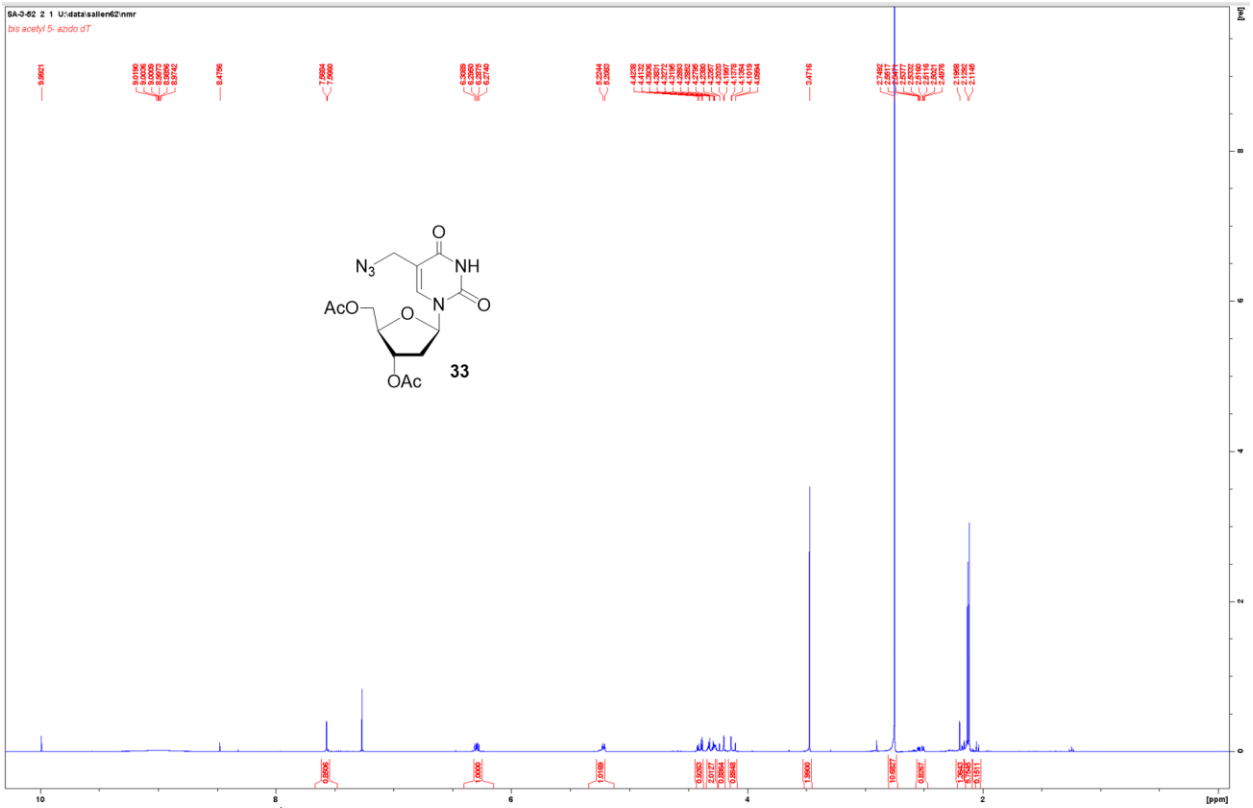
Appendix Figure 43. ^{1}H NMR Spectrum of **30**.



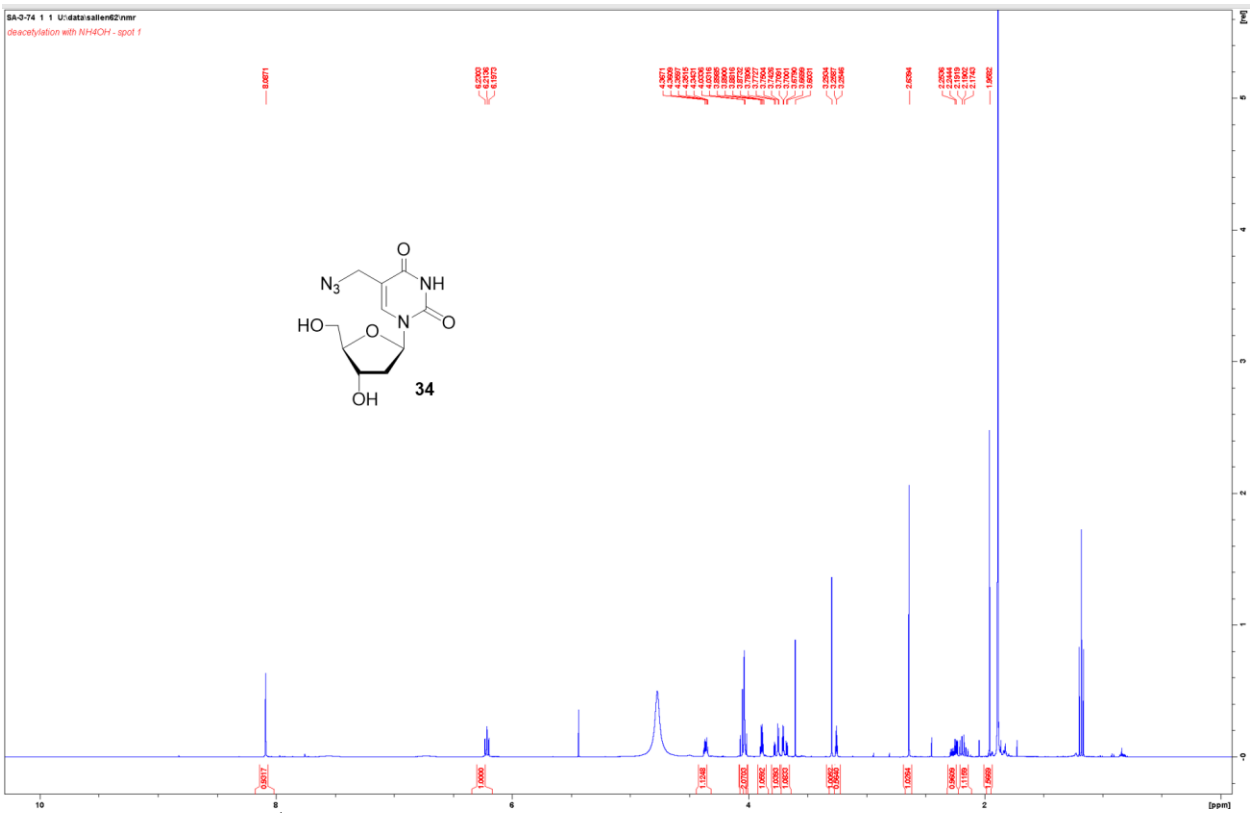
Appendix Figure 44. ^{1}H NMR Spectrum of **31**.



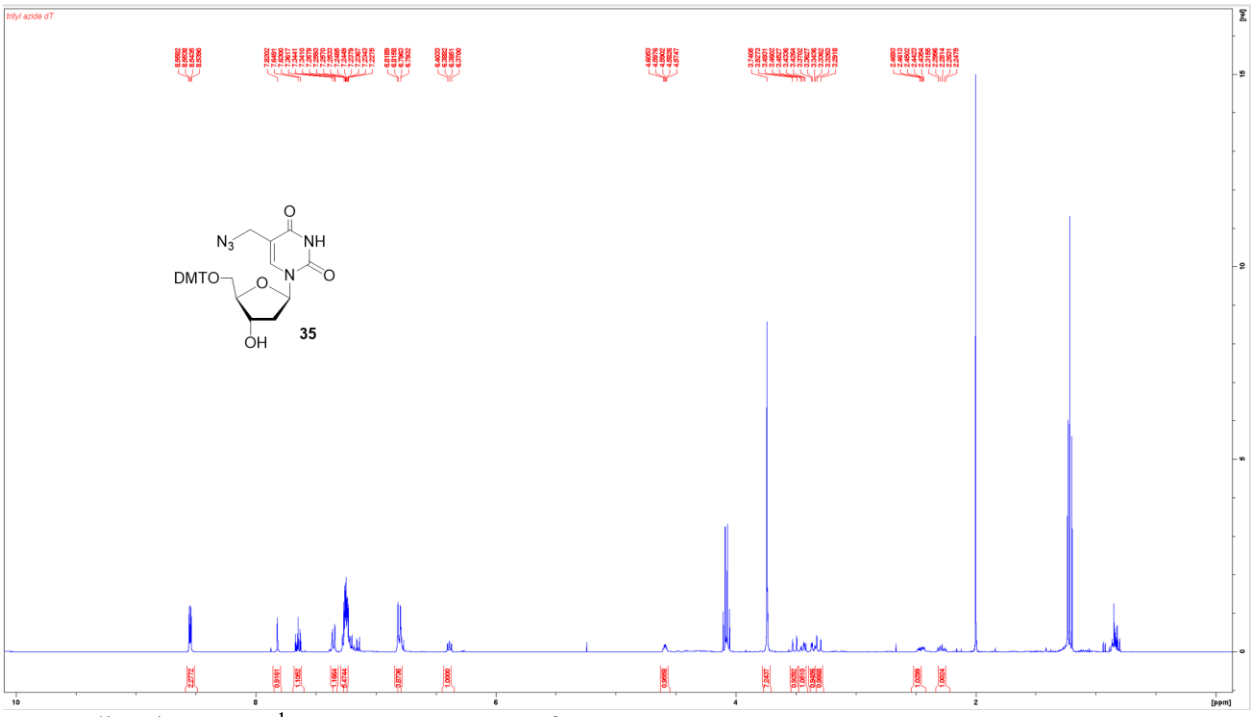
Appendix Figure 45. ^{1}H NMR Spectrum of **32**.



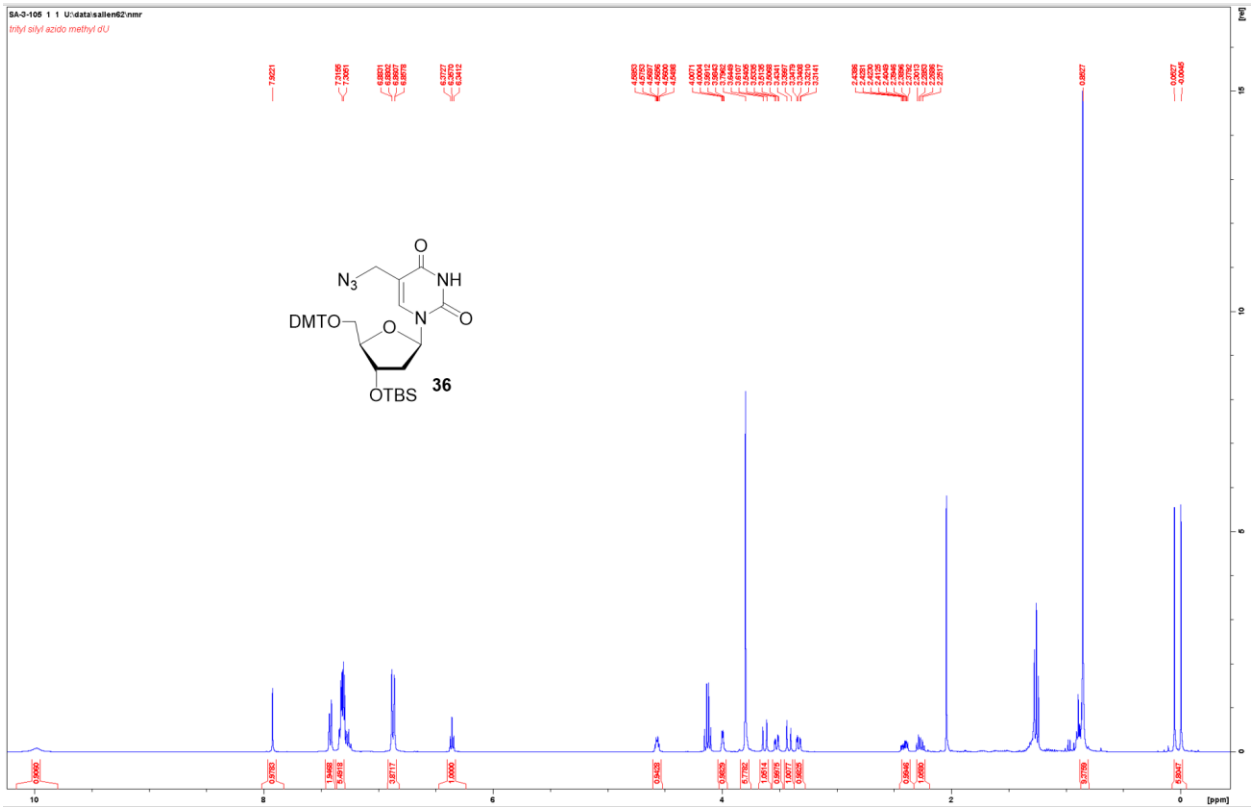
Appendix Figure 46. ^{1}H NMR Spectrum of **33**.



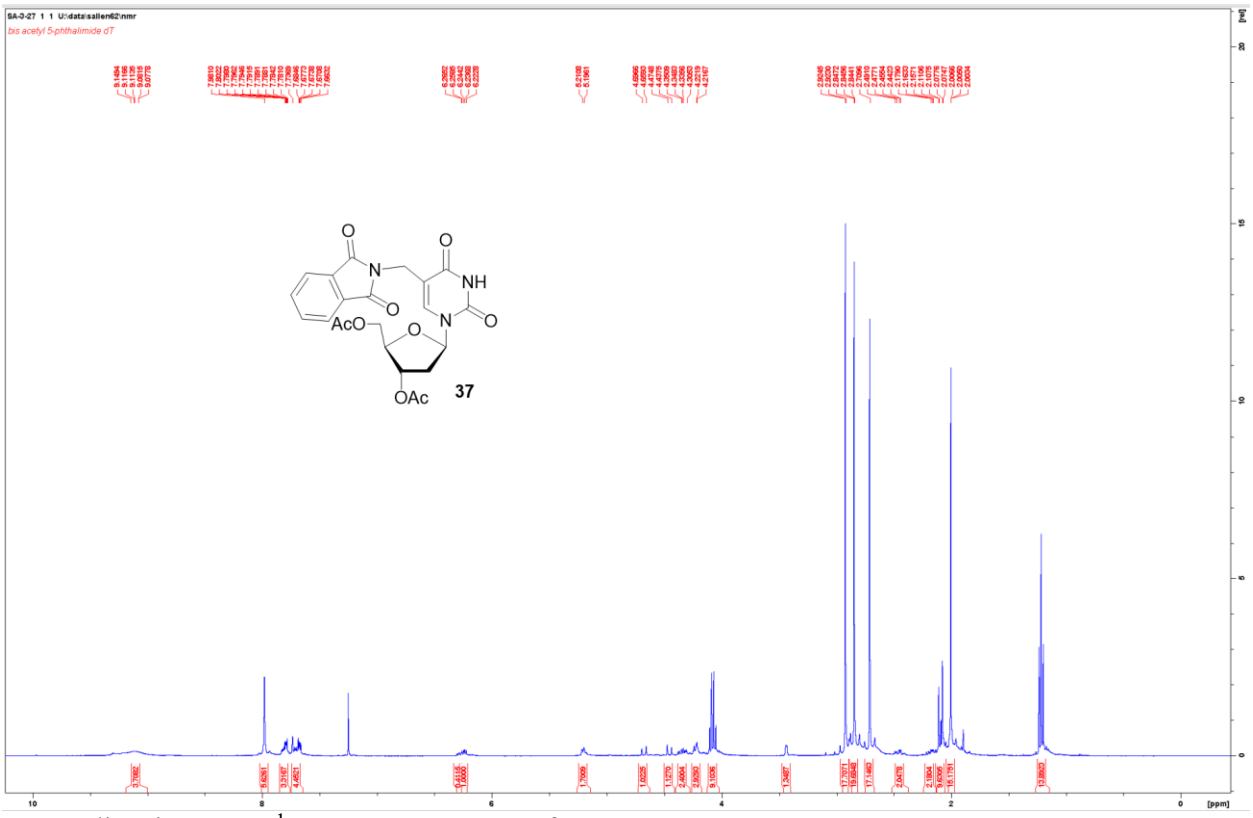
Appendix Figure 47. ^{1}H NMR Spectrum of **34**.



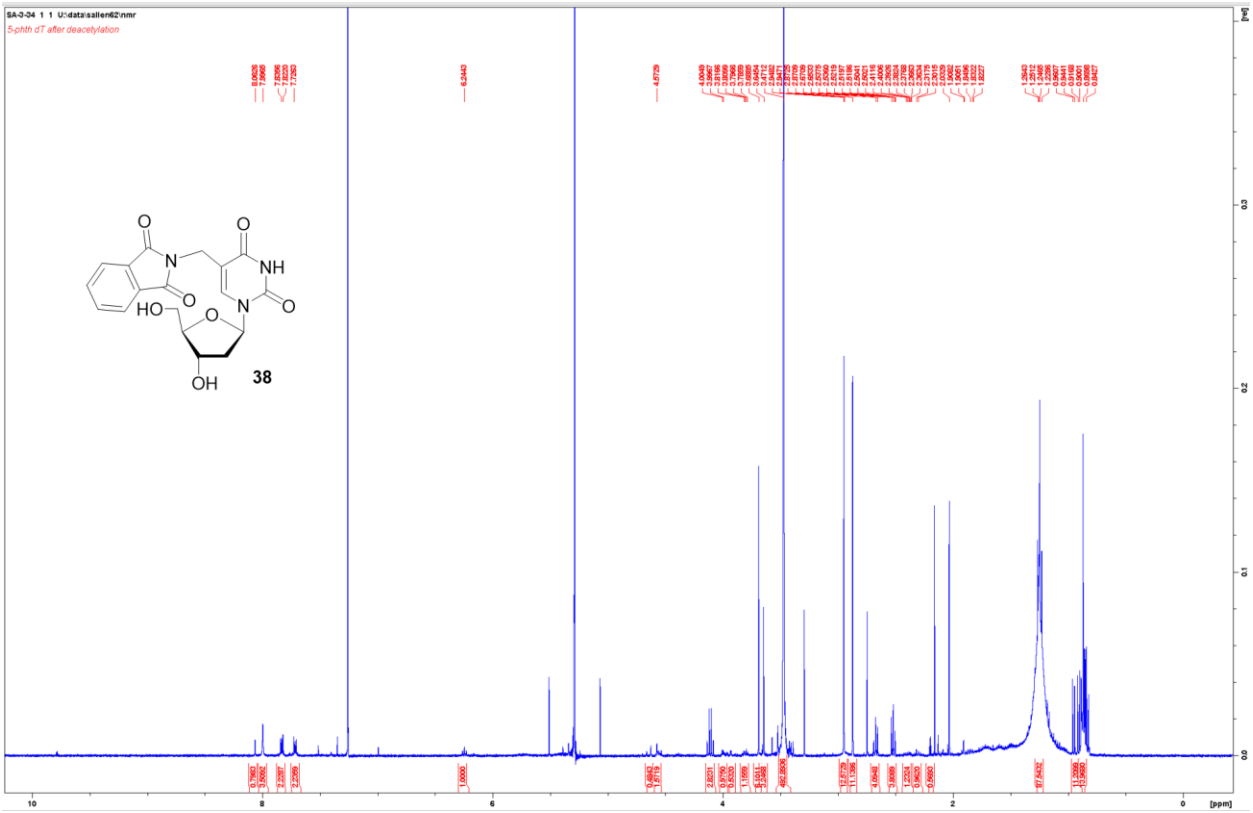
Appendix Figure 48. ^{1}H NMR Spectrum of **35**.



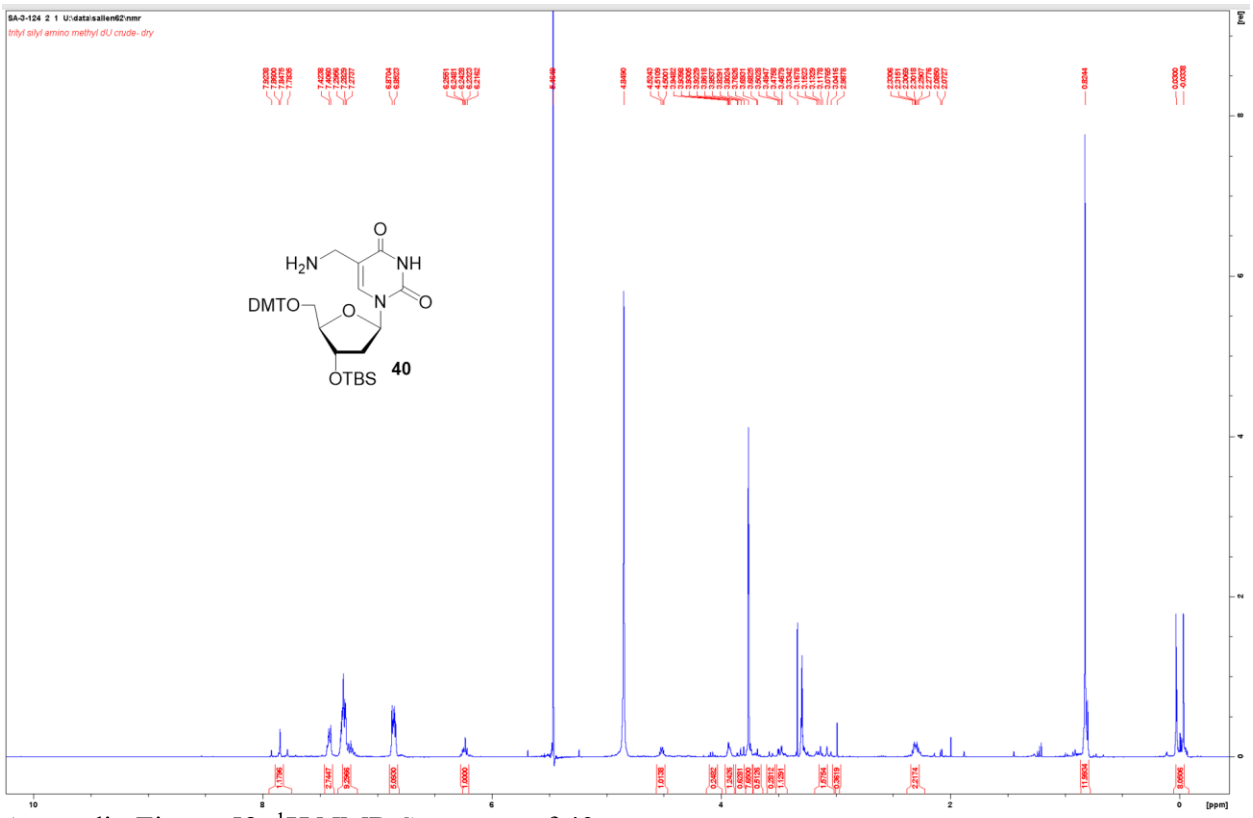
Appendix Figure 49. ^{1}H NMR Spectrum of **36**.



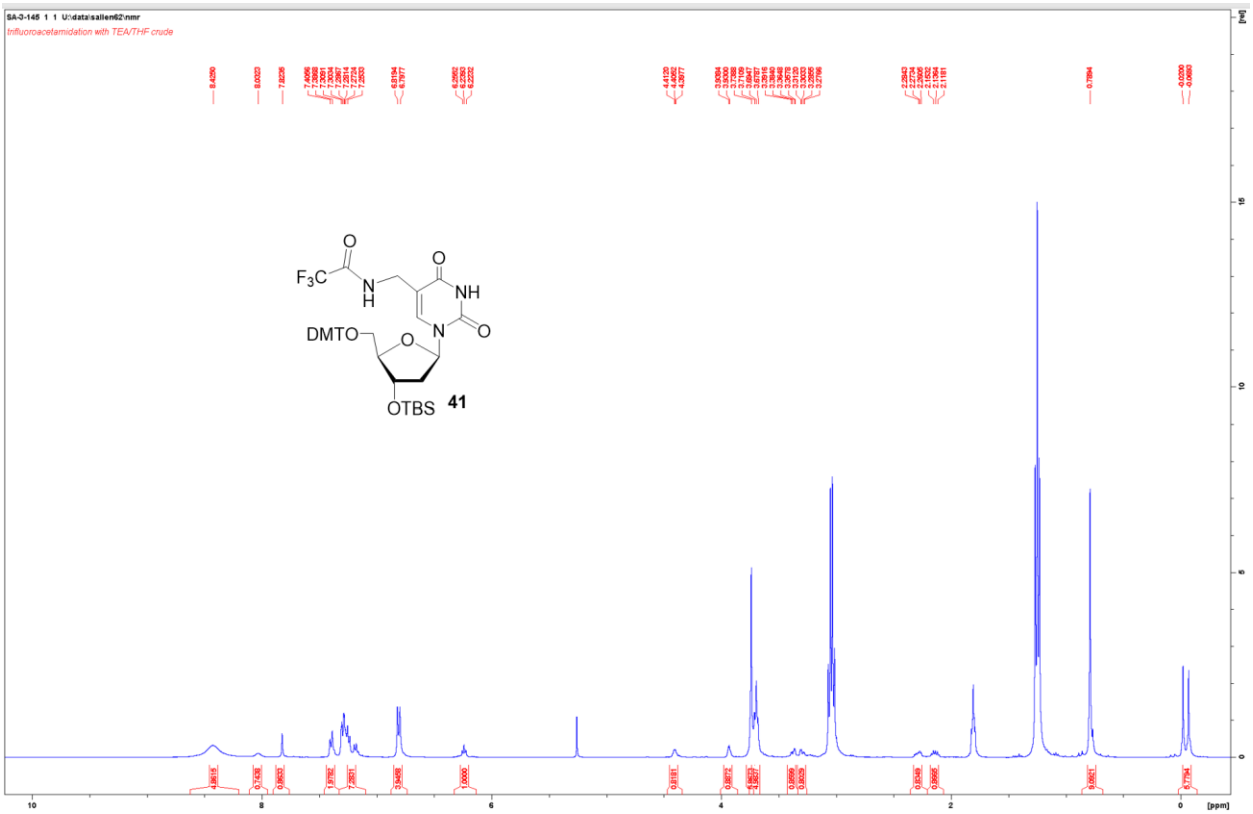
Appendix Figure 50. ^{1}H NMR Spectrum of **37**.



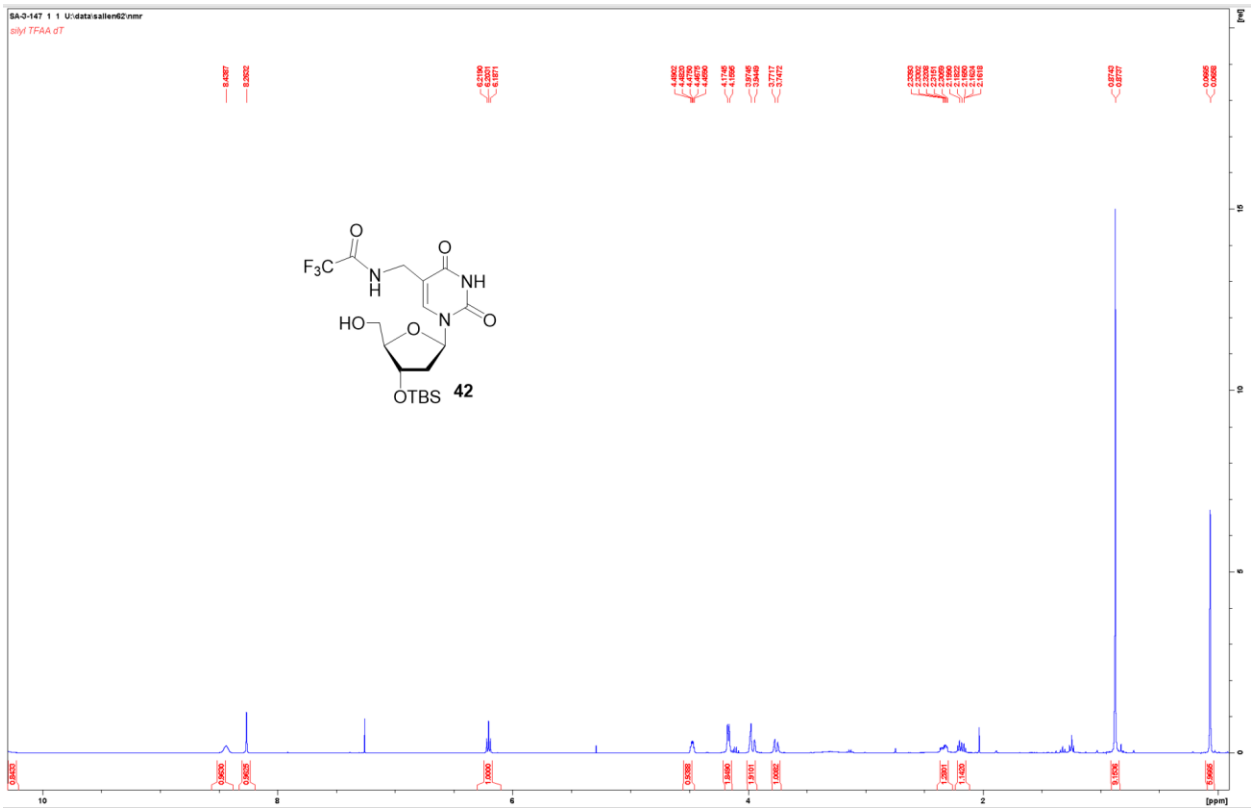
Appendix Figure 51. ^{1}H NMR Spectrum of **38**.



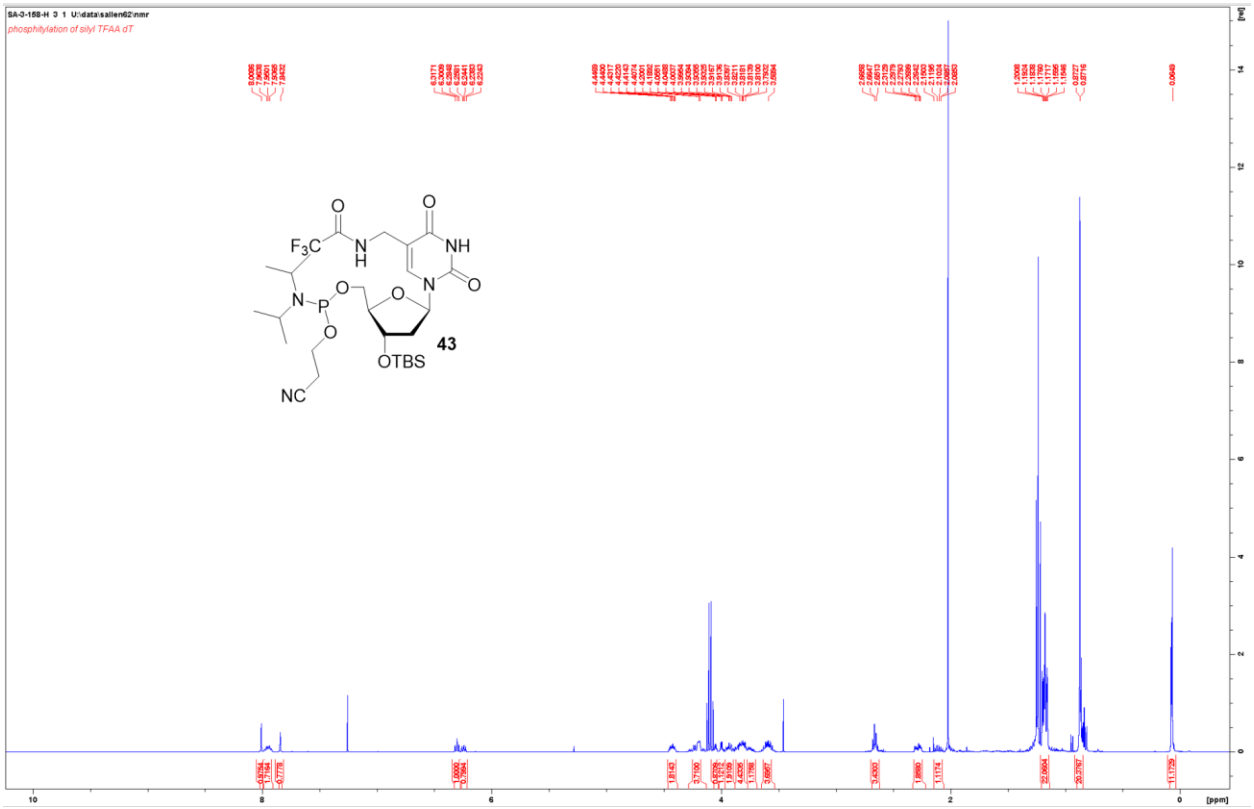
Appendix Figure 52. ^{1}H NMR Spectrum of **40**.



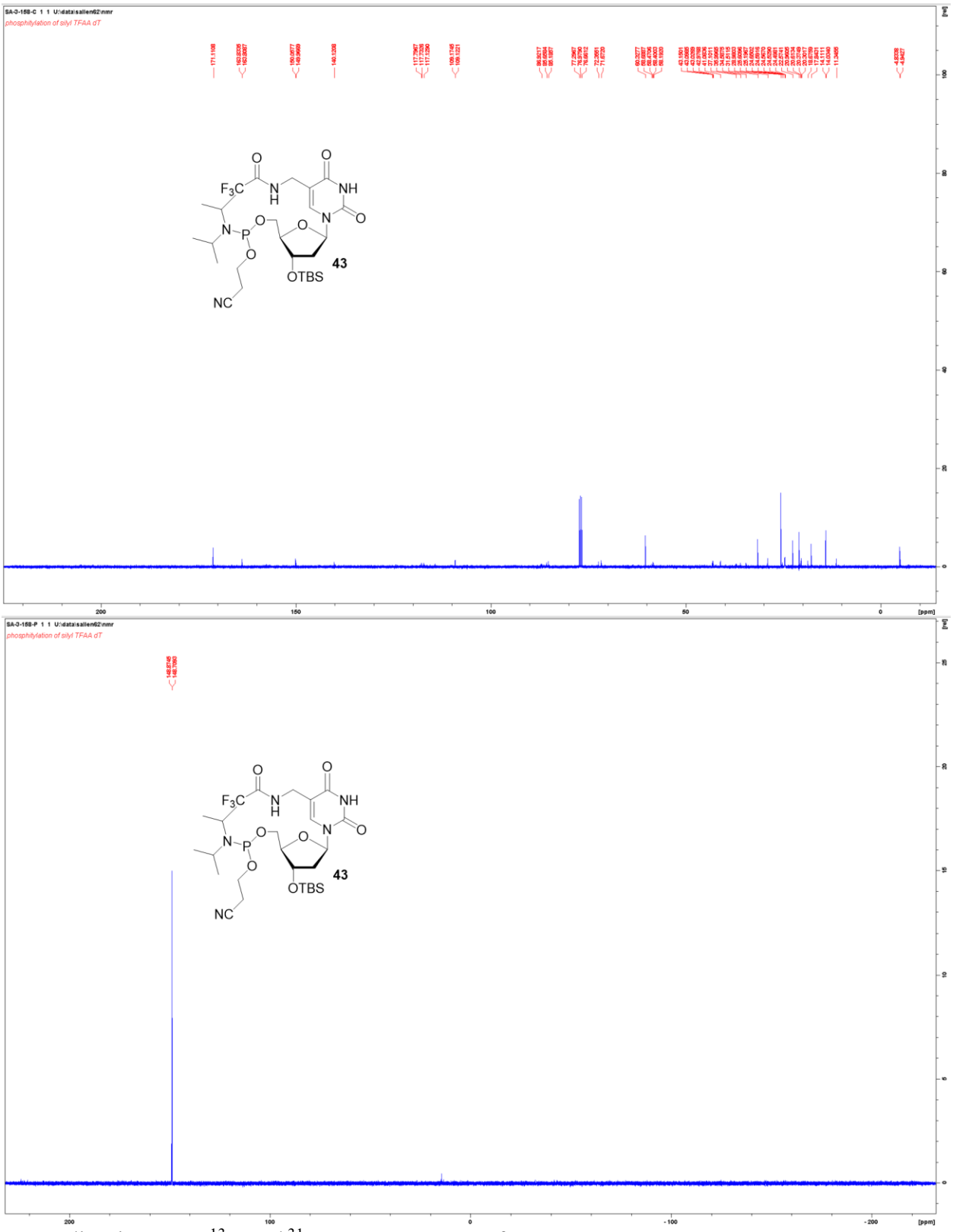
Appendix Figure 53. ^{1}H NMR Spectrum of **41**.



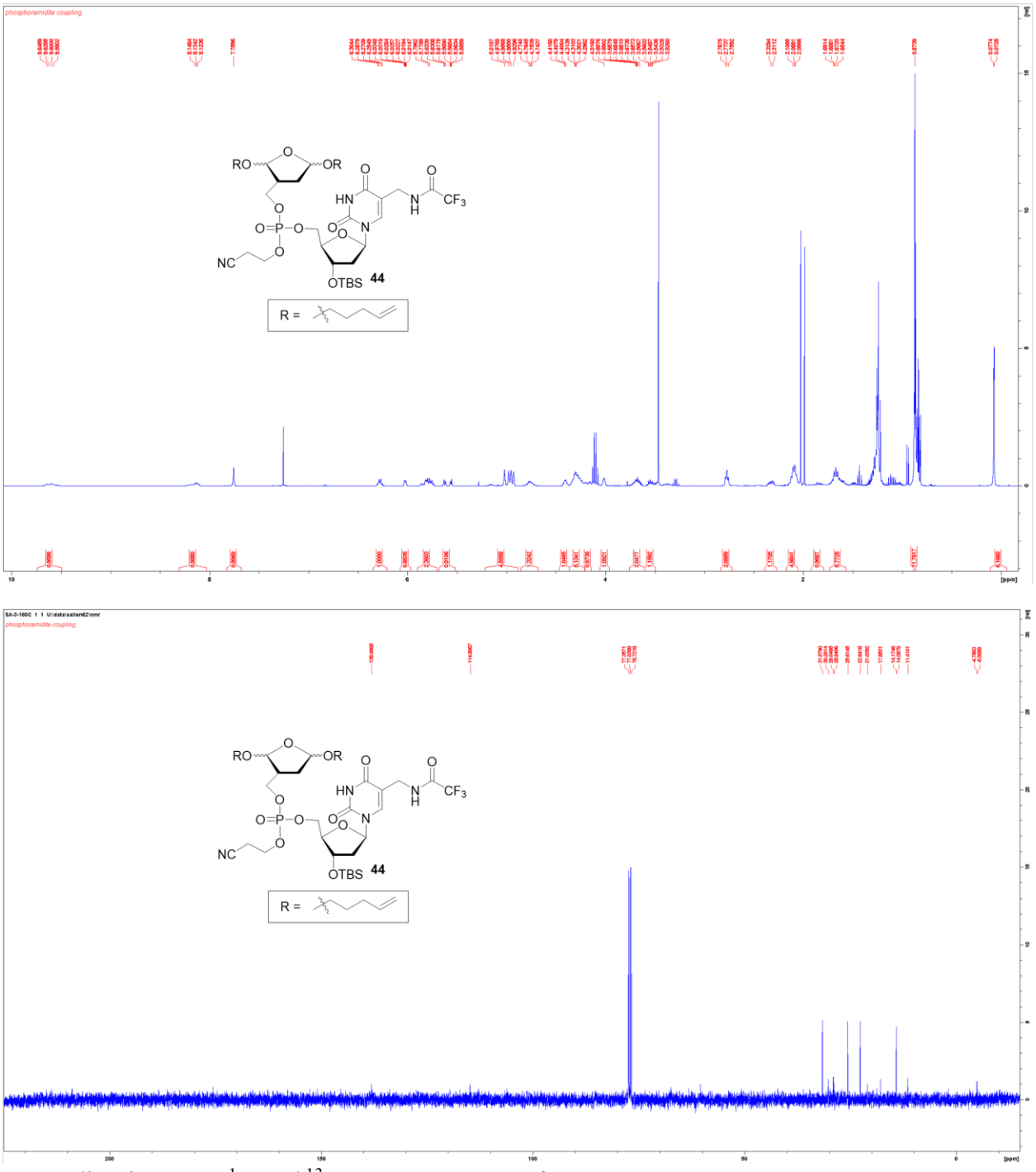
Appendix Figure 54. ^{1}H NMR Spectrum of **42.**



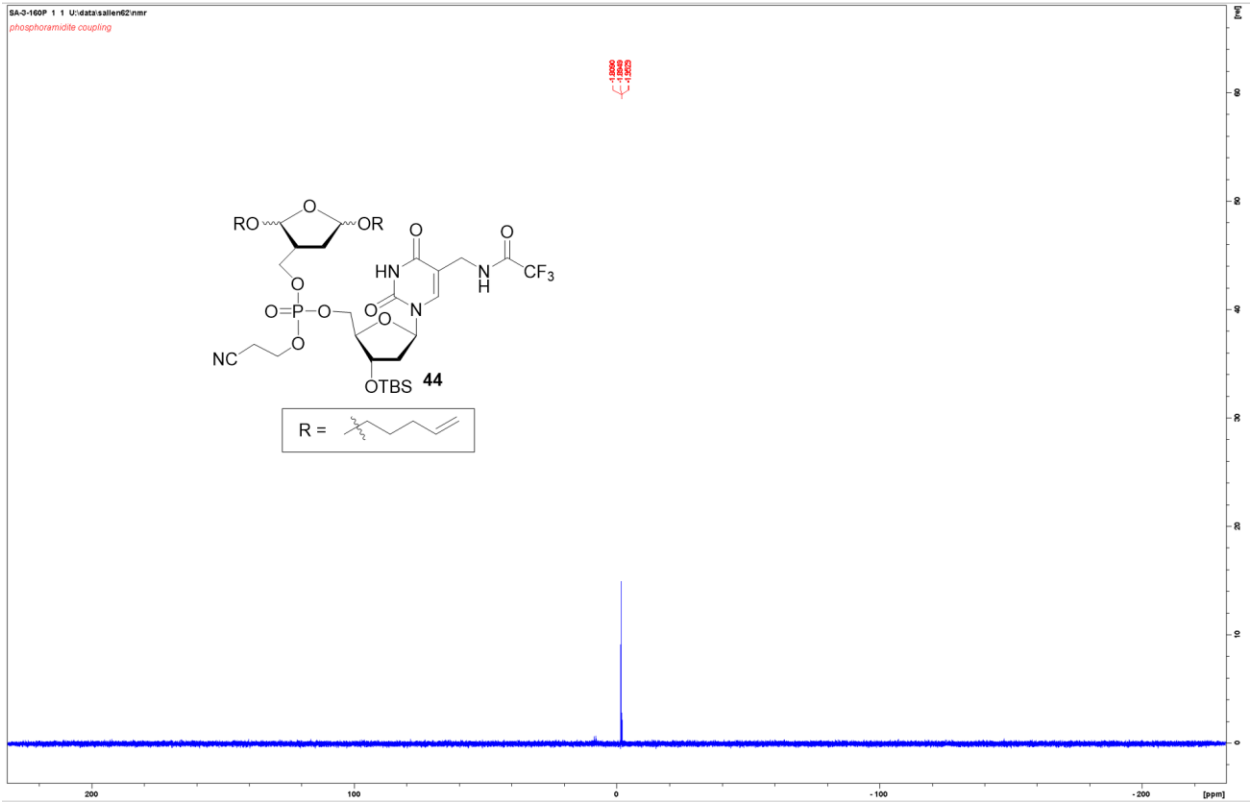
Appendix Figure 55. ^{1}H NMR Spectrum of **43.**



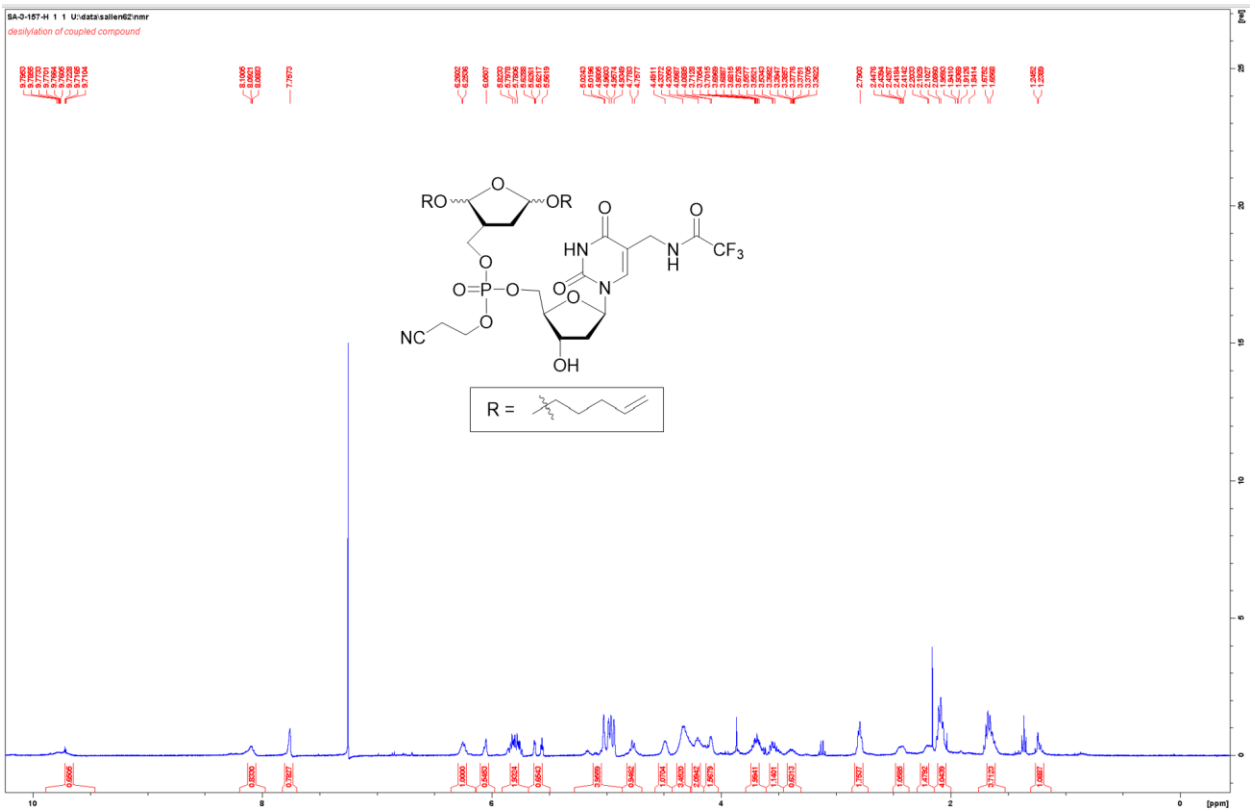
Appendix Figure 56. ^{13}C and ^{31}P NMR Spectrum of **43.**



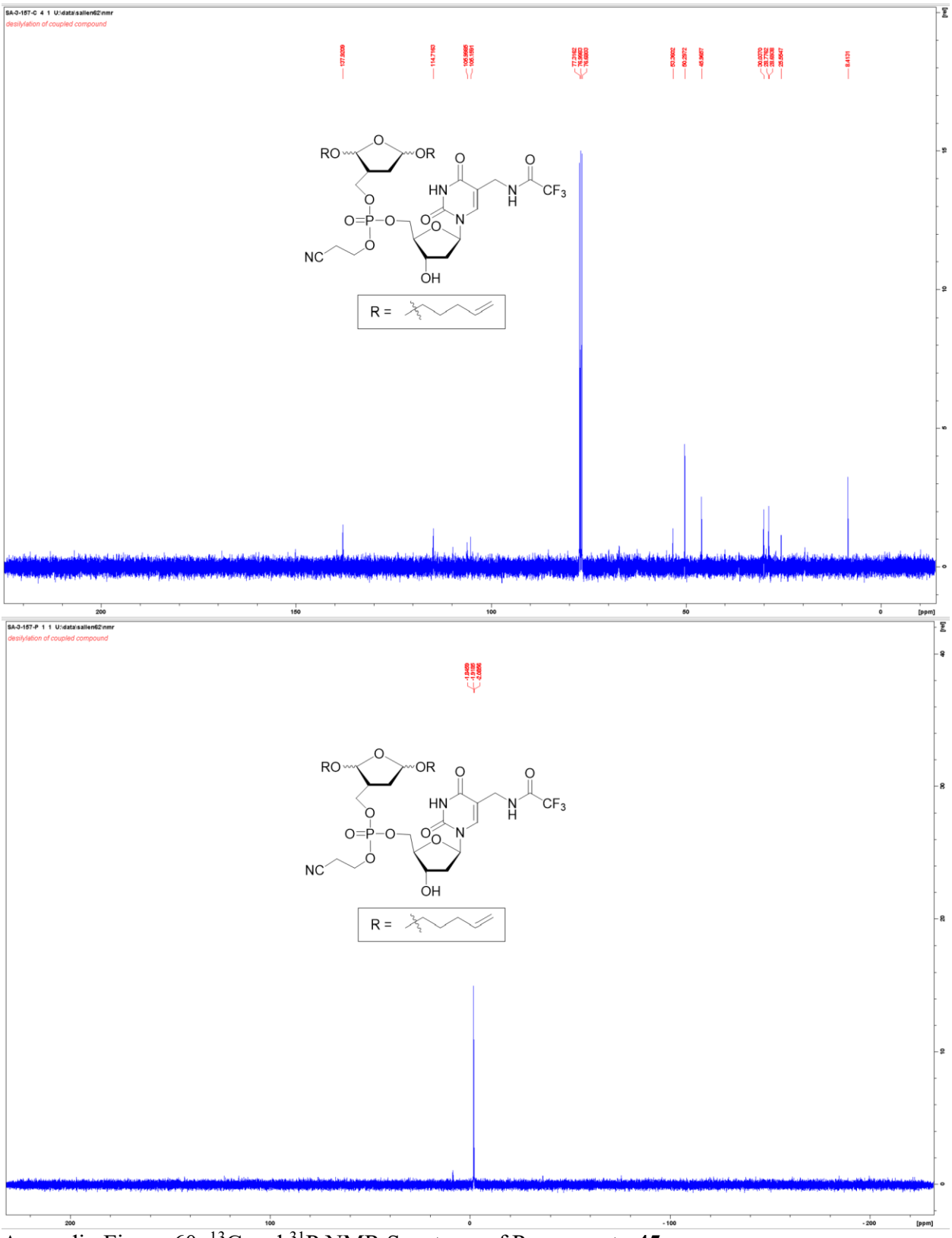
Appendix Figure 57. ^{1}H and ^{13}C NMR Spectrum of **44.**



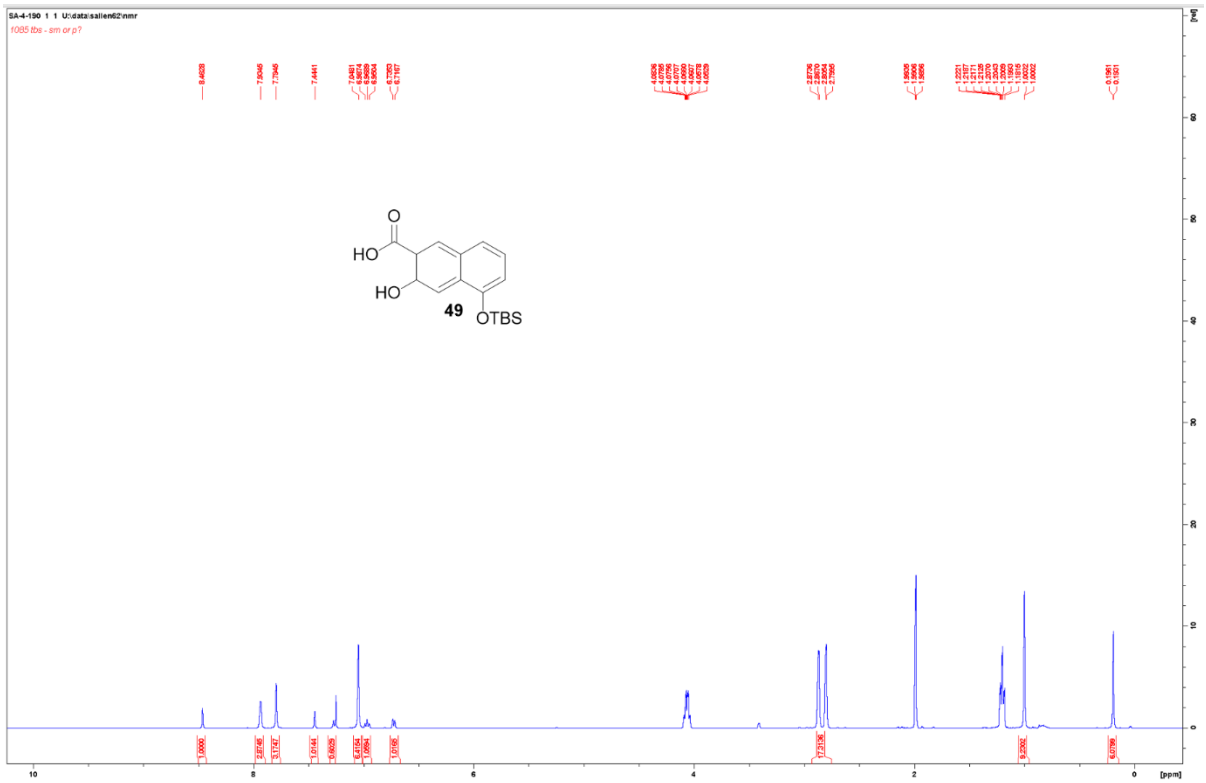
Appendix Figure 58. ^{31}P NMR Spectrum of **44.**



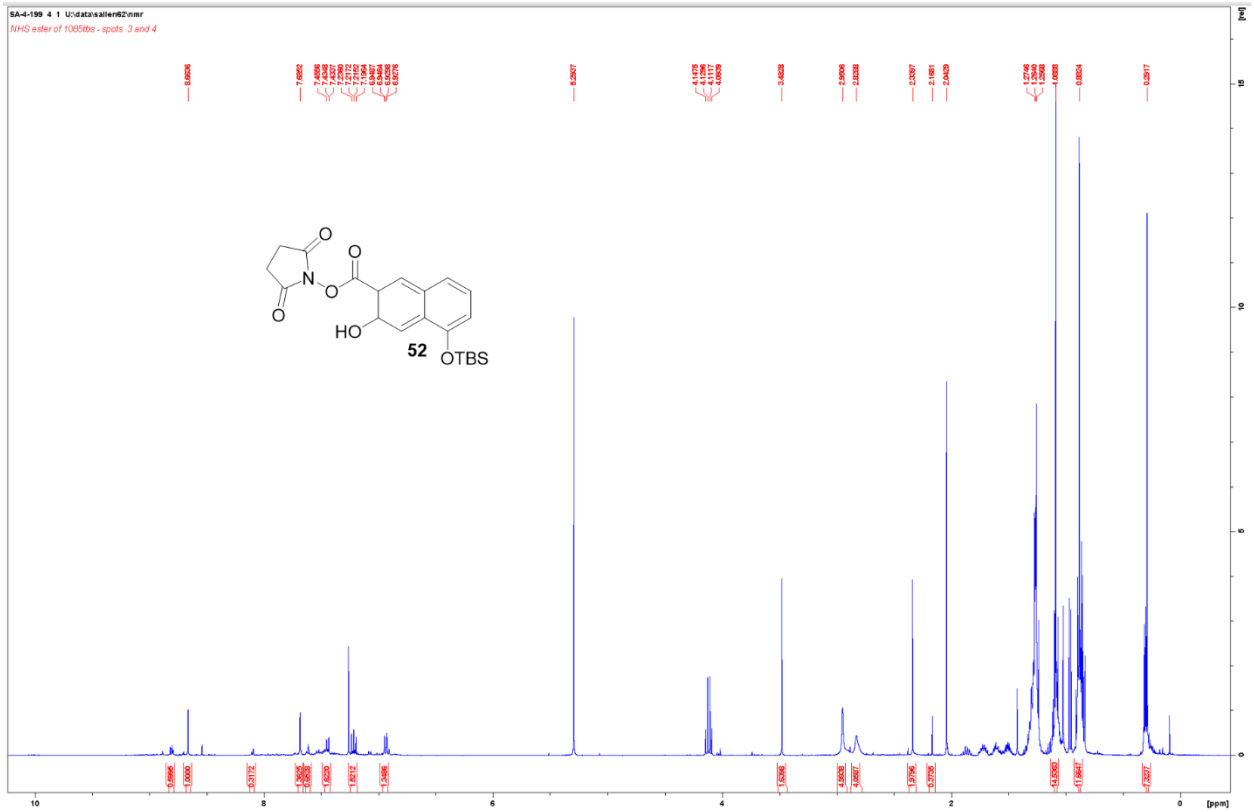
Appendix Figure 59. ^{1}H NMR Spectrum of Precursor to **45.**



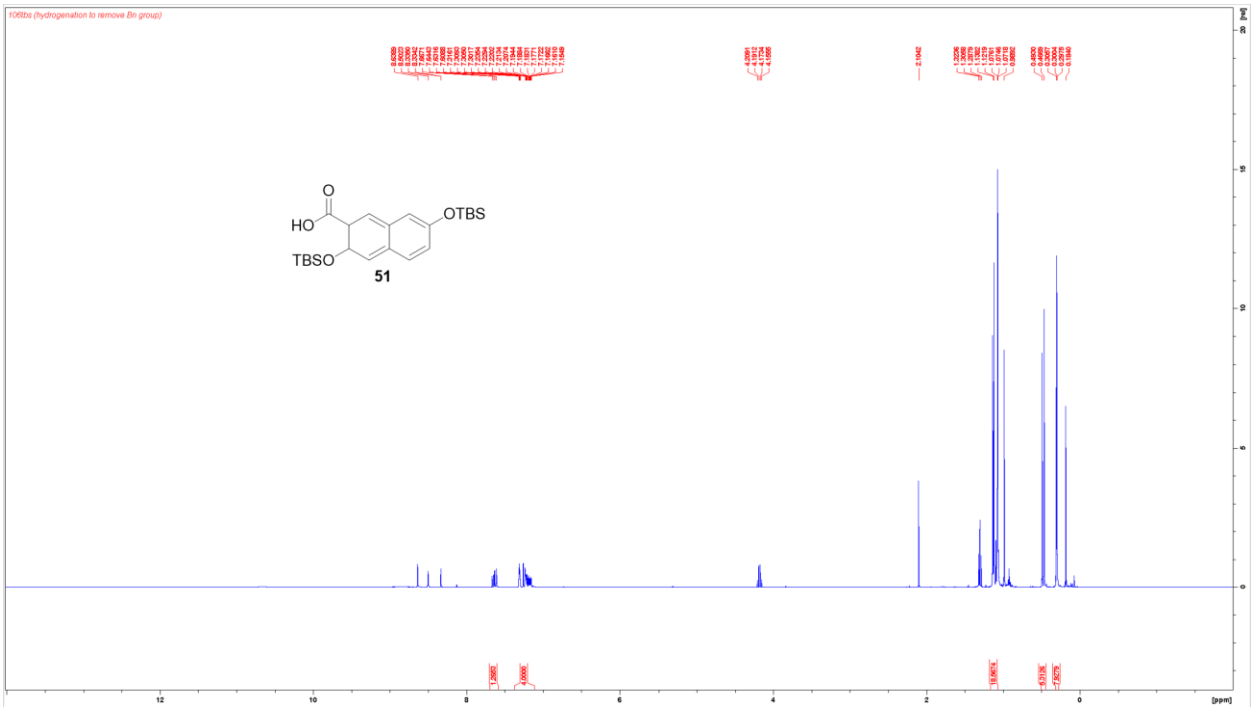
Appendix Figure 60. ^{13}C and ^{31}P NMR Spectrum of Precursor to **45.**



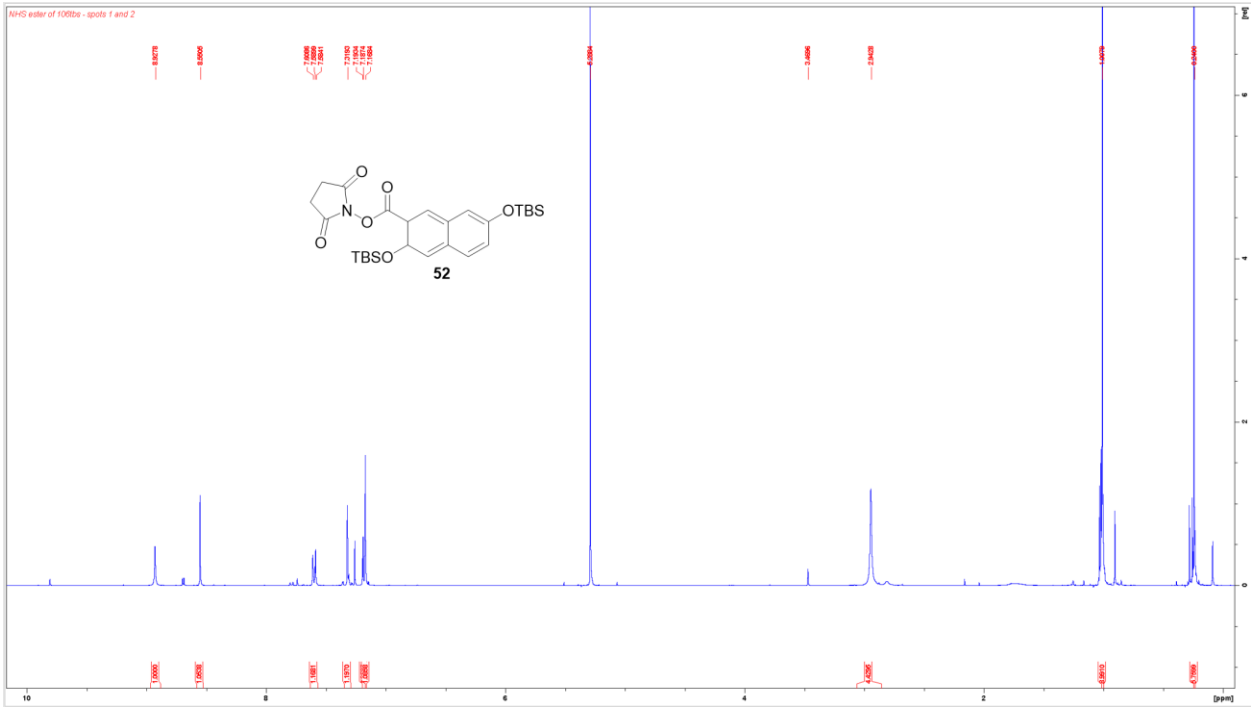
Appendix Figure 61. ^{1}H NMR Spectrum of **49.**



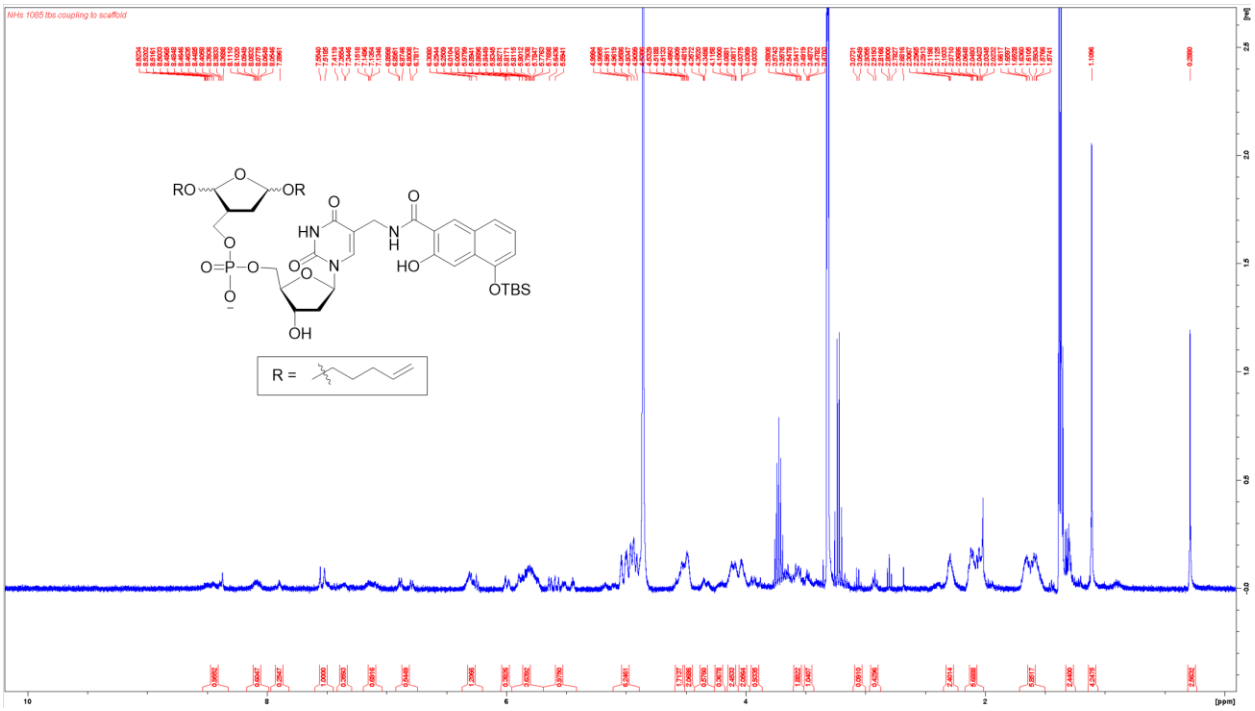
Appendix Figure 62. ^{1}H NMR Spectrum of **50.**



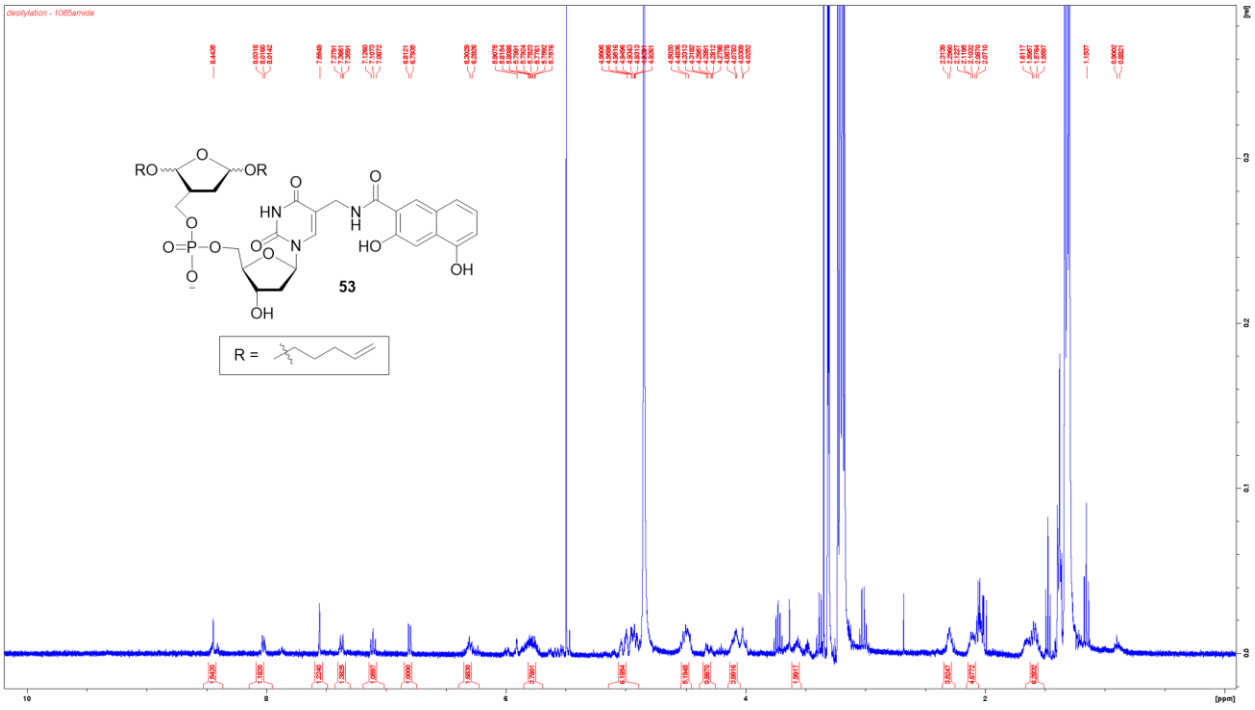
Appendix Figure 63. ^{1}H NMR Spectrum of **51.**



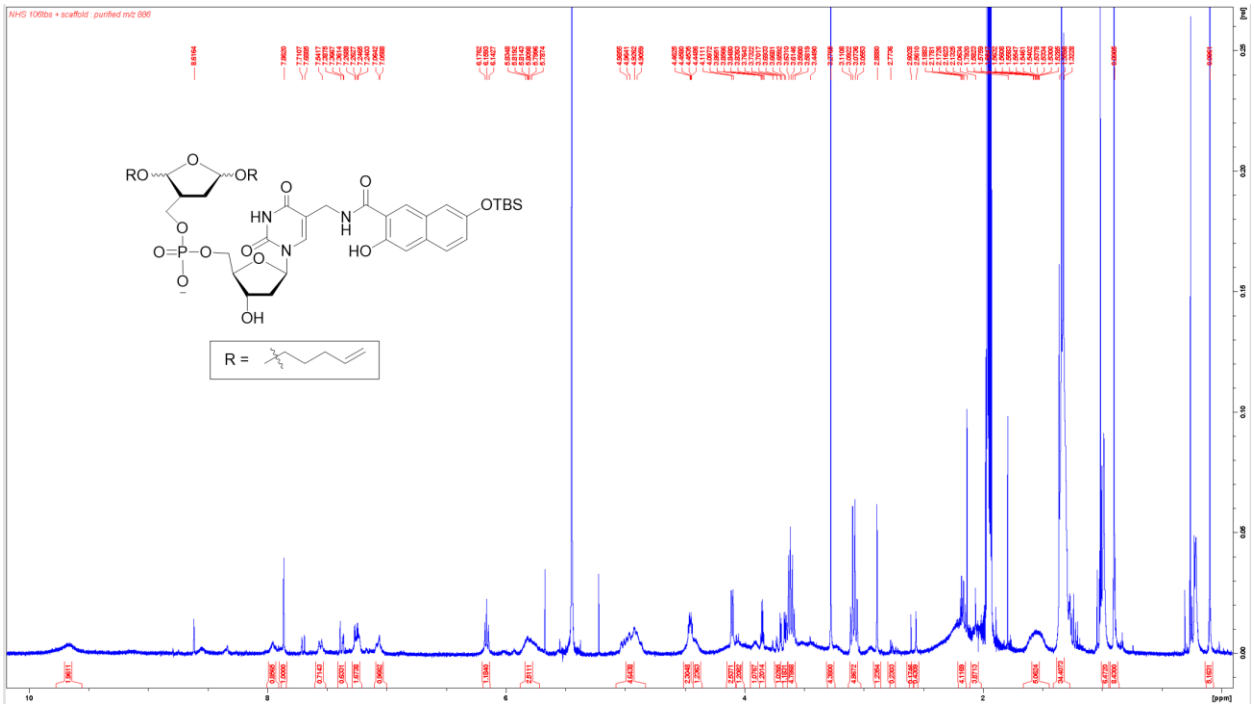
Appendix Figure 64. ^{1}H NMR Spectrum of **52.**



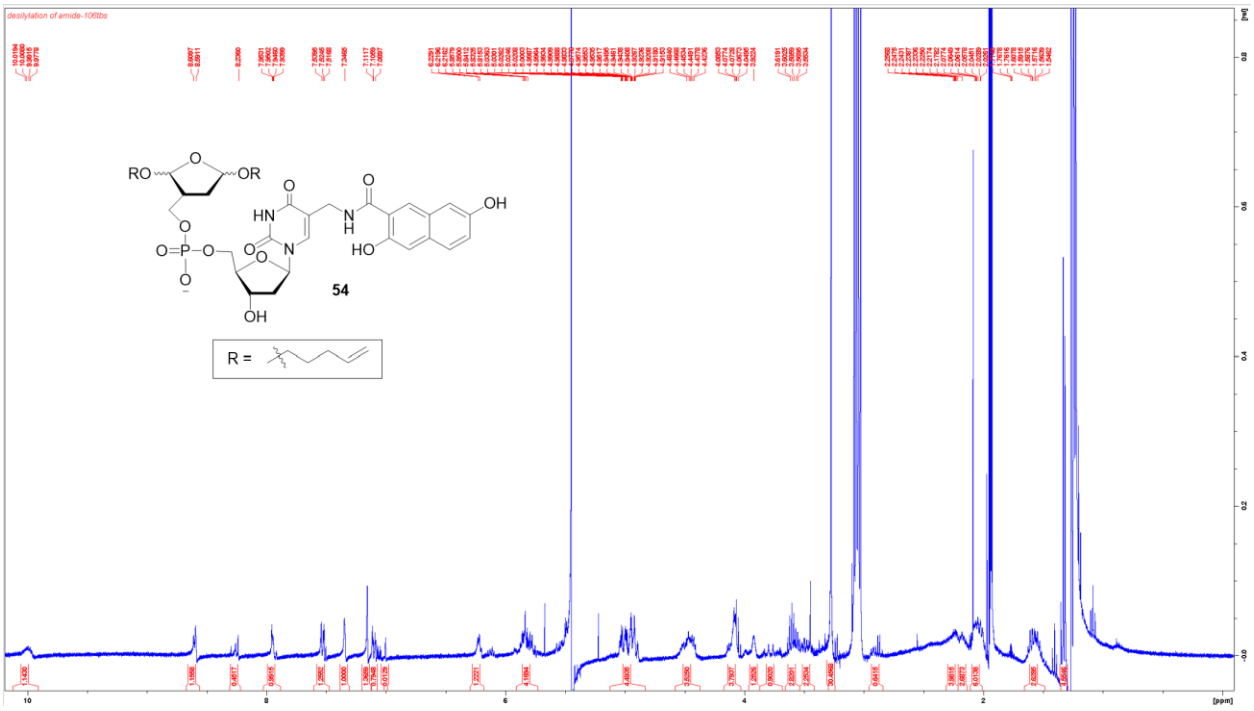
Appendix Figure 65. ^{1}H NMR Spectrum of Precursor to **53.**



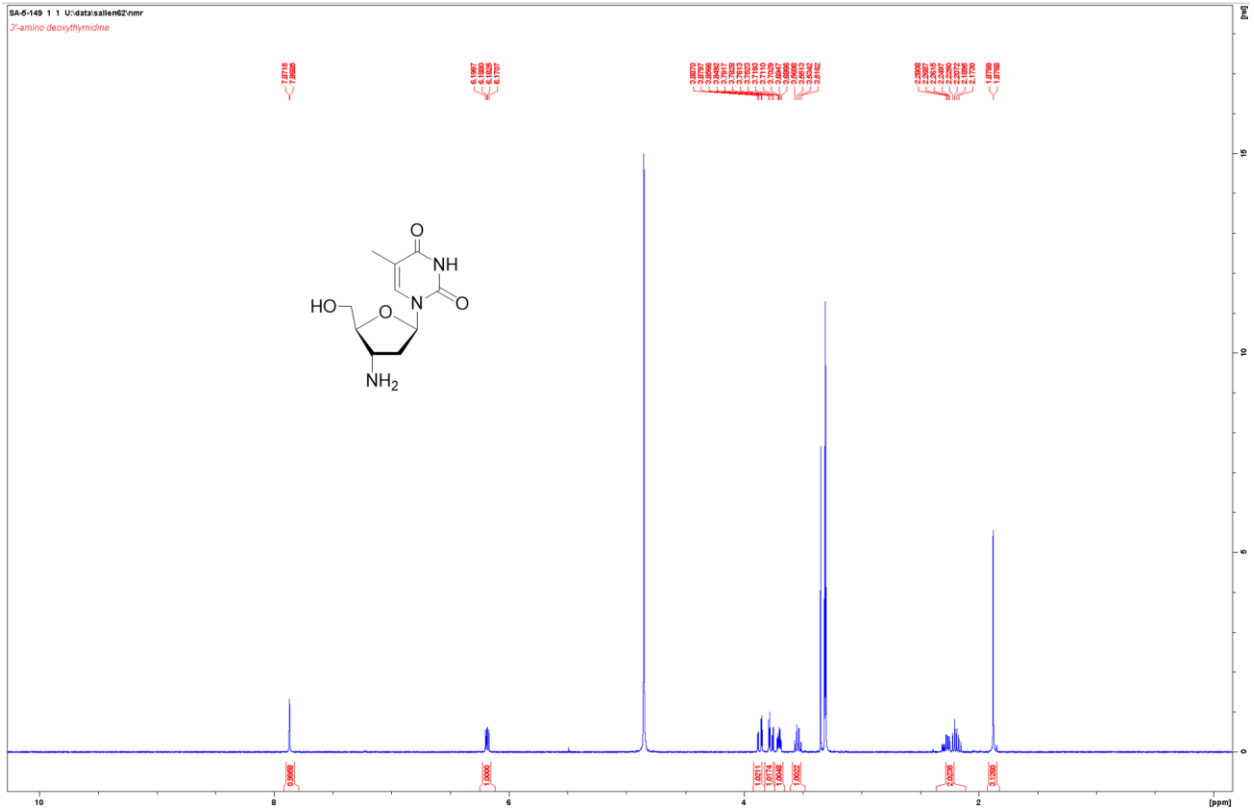
Appendix Figure 66. ^{1}H NMR Spectrum of **53.**



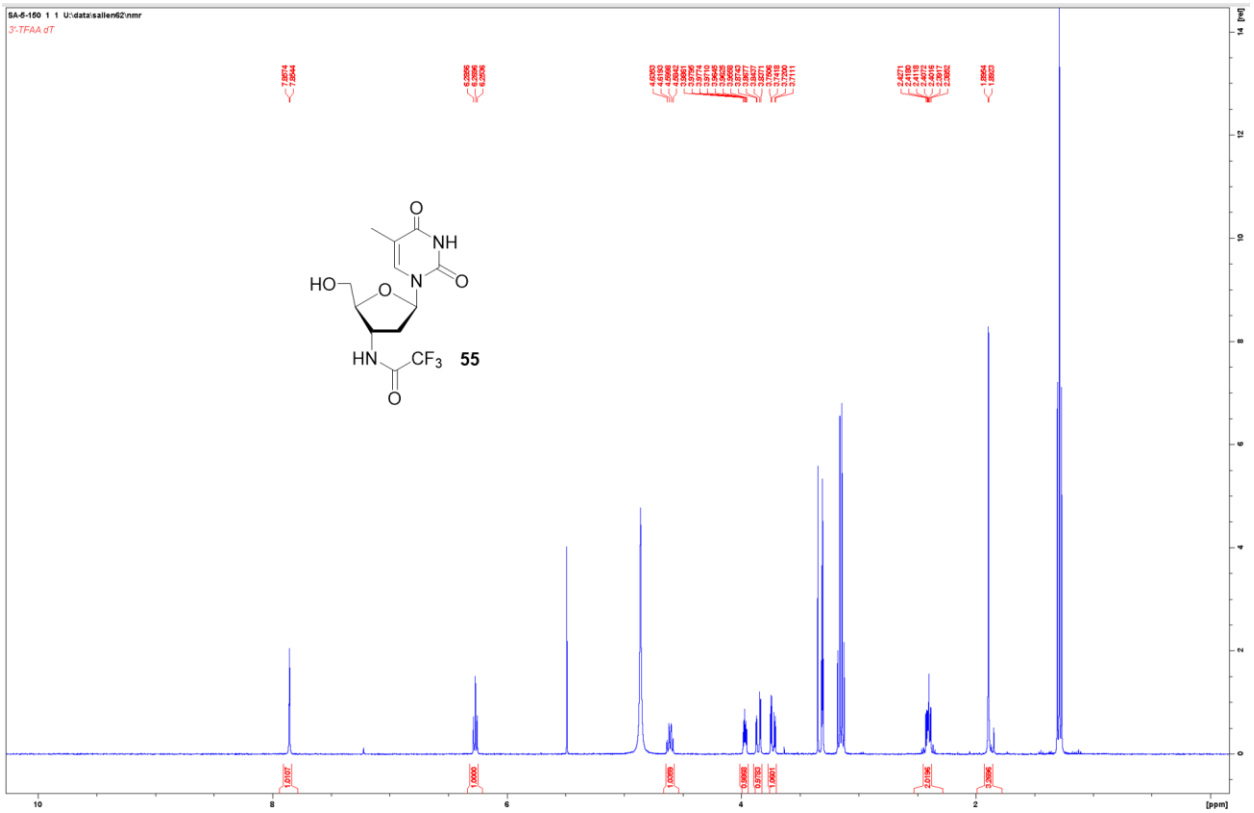
Appendix Figure 67. ^{1}H NMR Spectrum of Precursor to **54.**



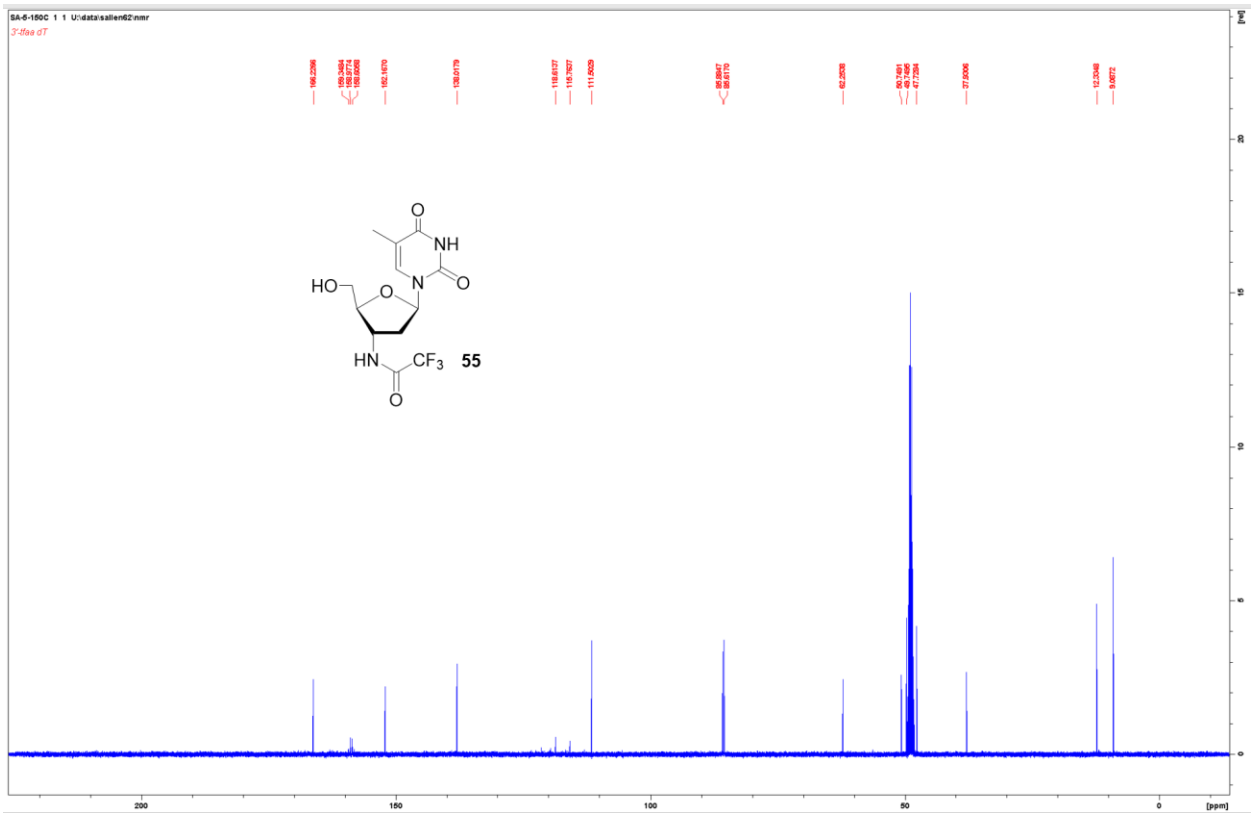
Appendix Figure 68. ^{1}H NMR Spectrum of **54.**



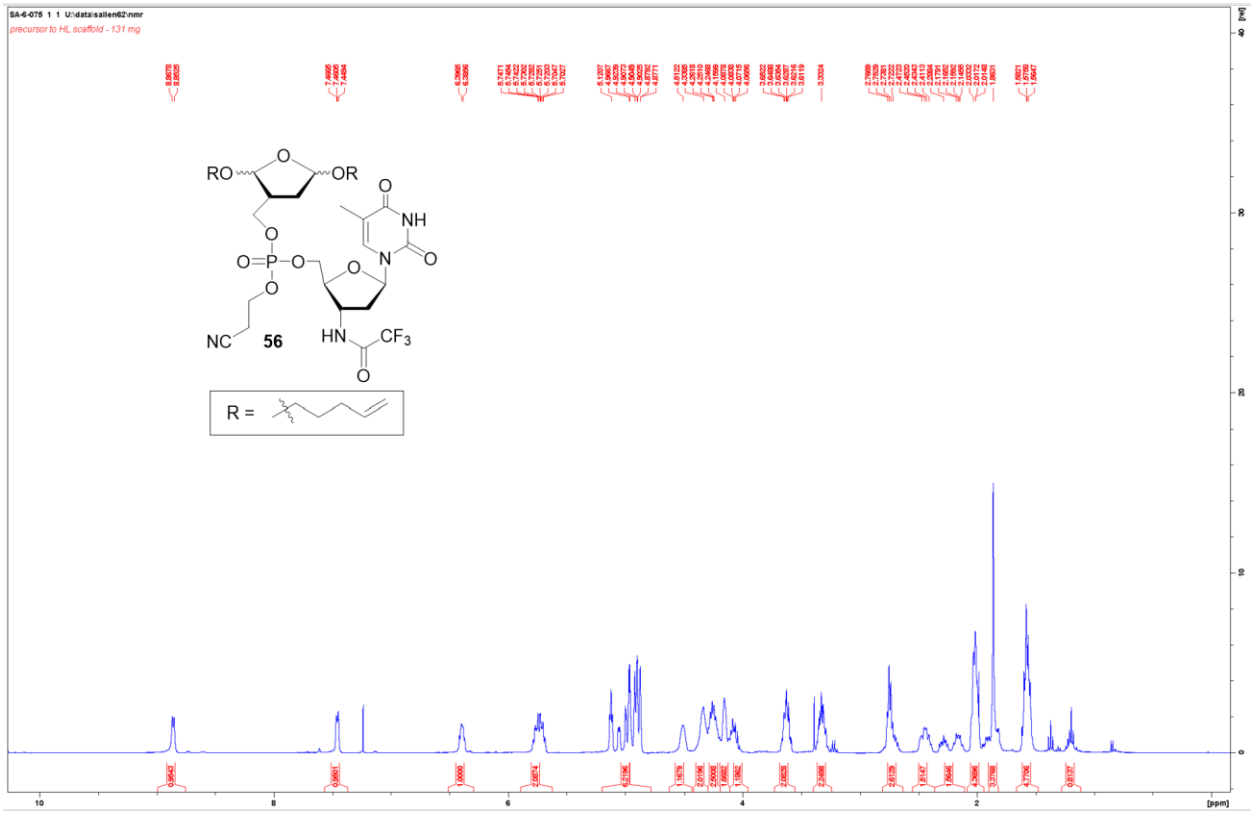
Appendix Figure 69. ^{1}H NMR Spectrum of Precursor to **55.**



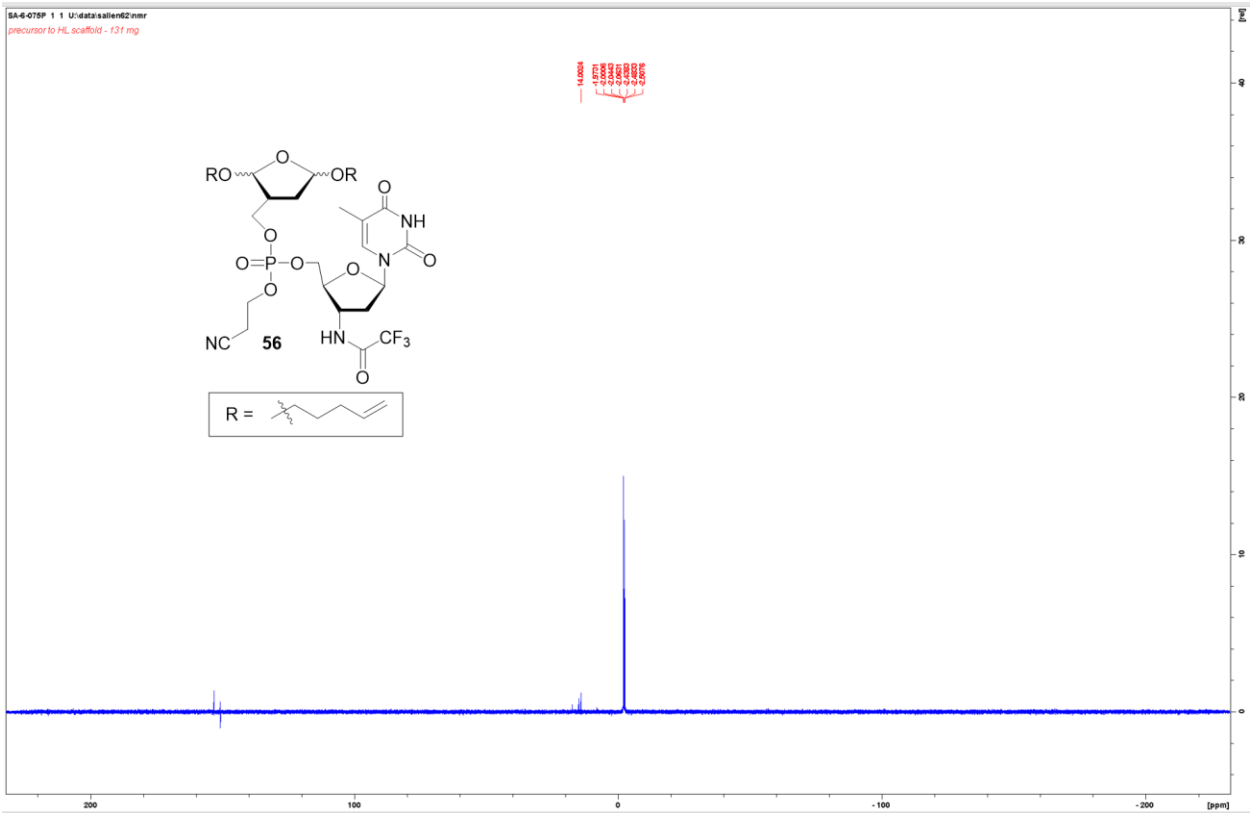
Appendix Figure 70. ^{1}H NMR Spectrum of **55.**



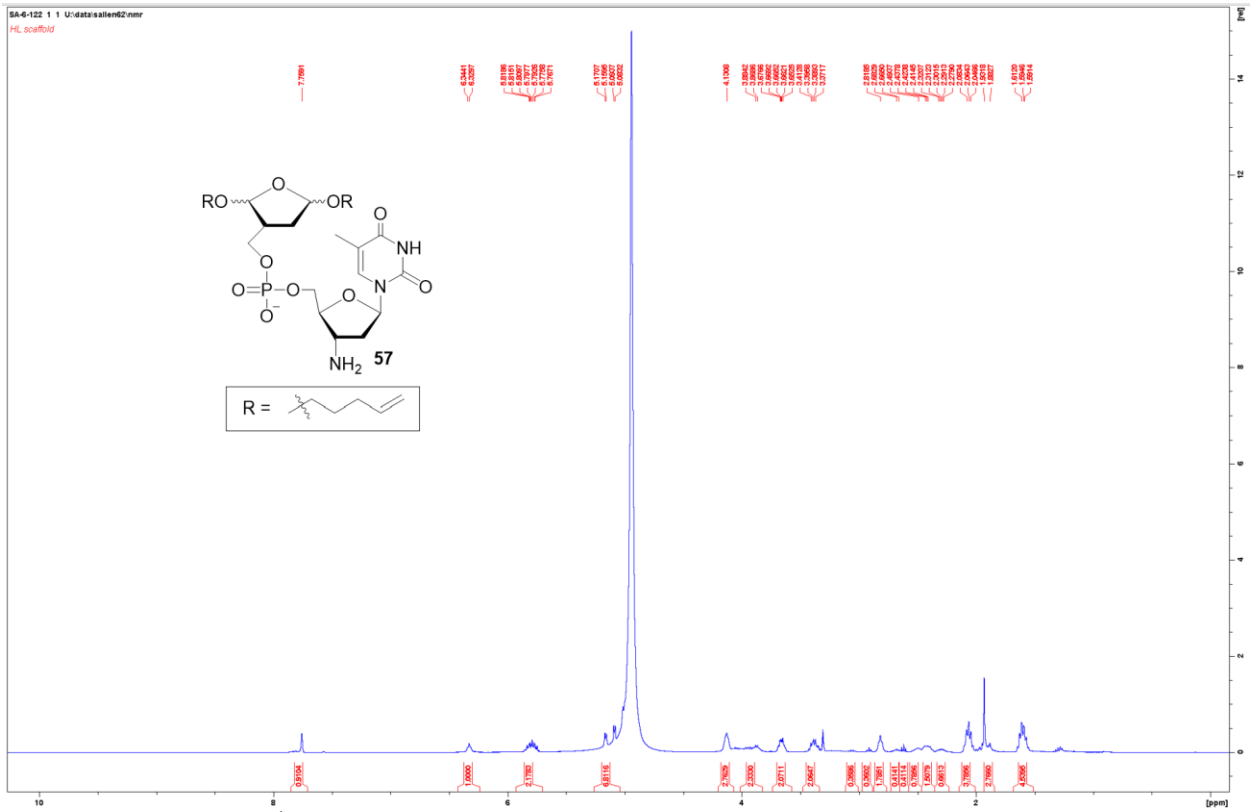
Appendix Figure 71. ^{13}C NMR Spectrum of **55.**



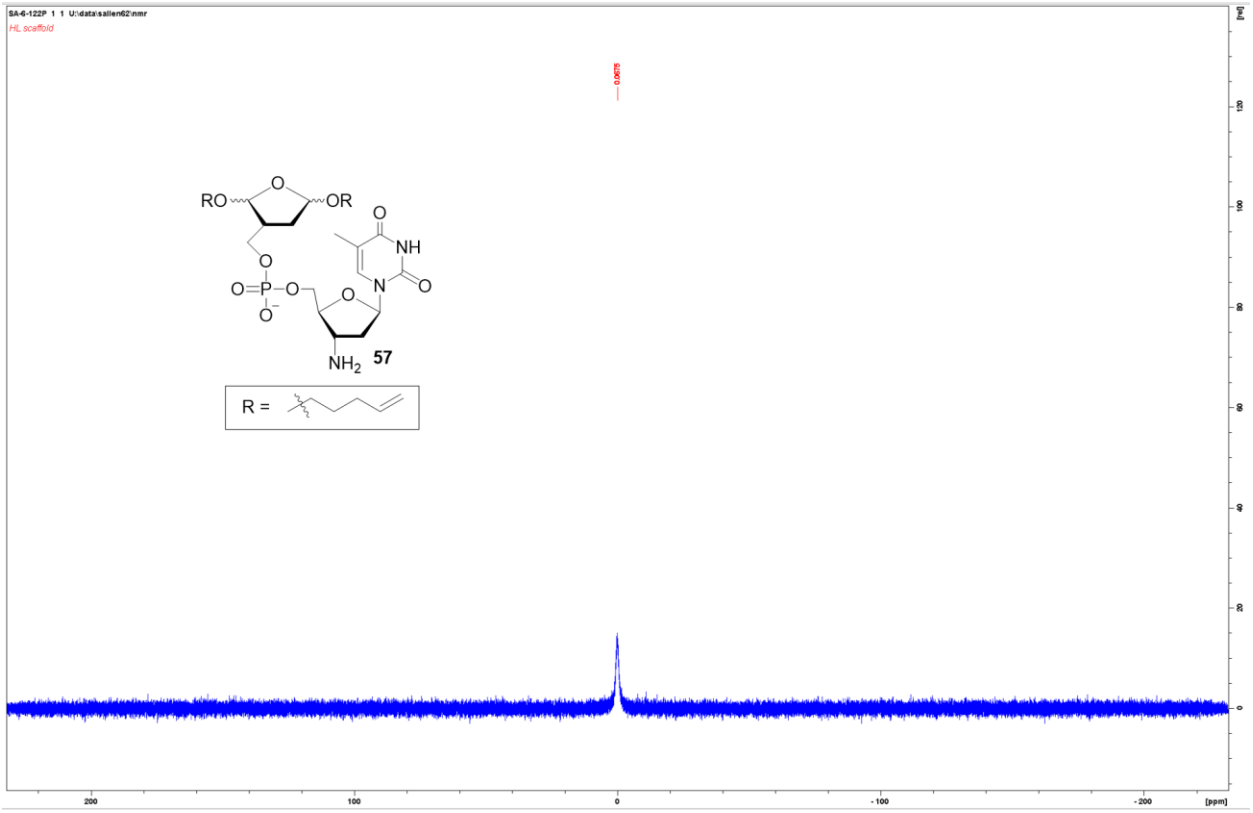
Appendix Figure 72. ^{1}H NMR Spectrum of **56.**



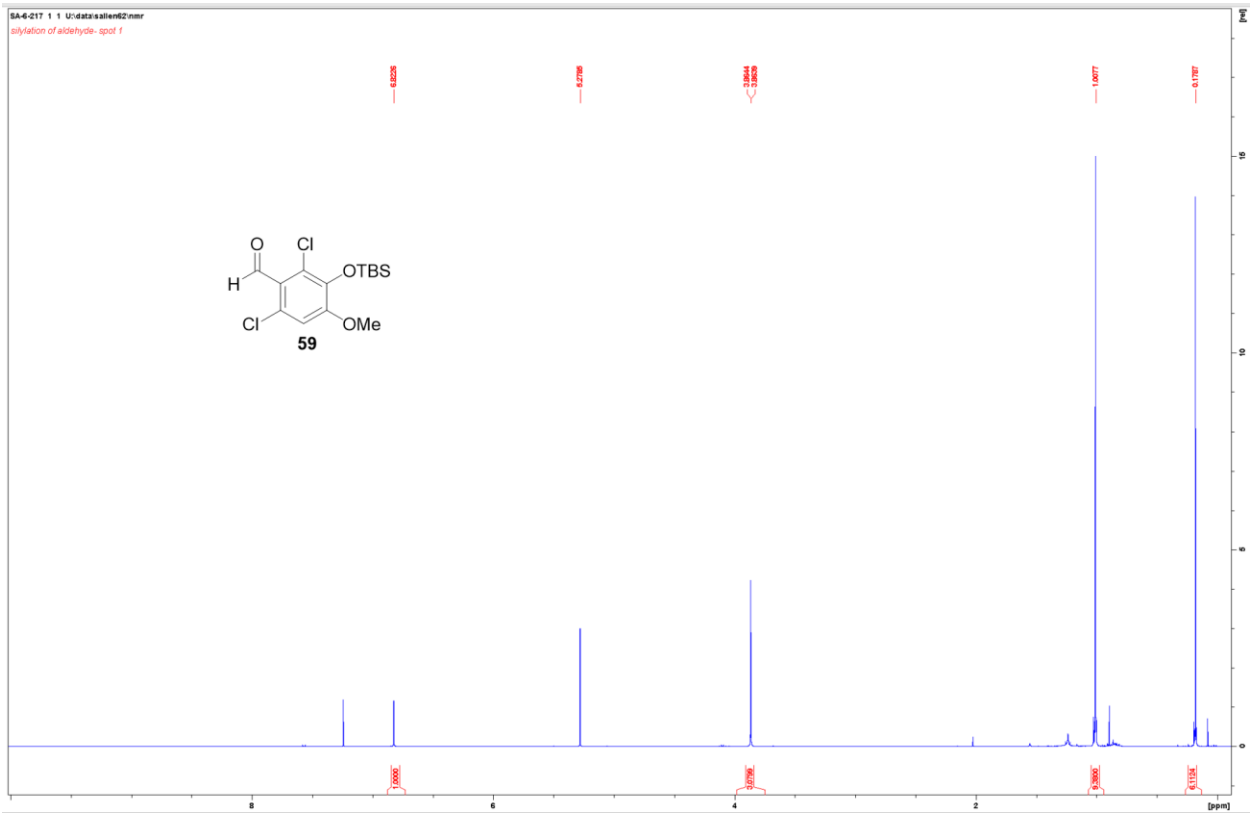
Appendix Figure 73. ^{31}P NMR Spectrum of **56.**



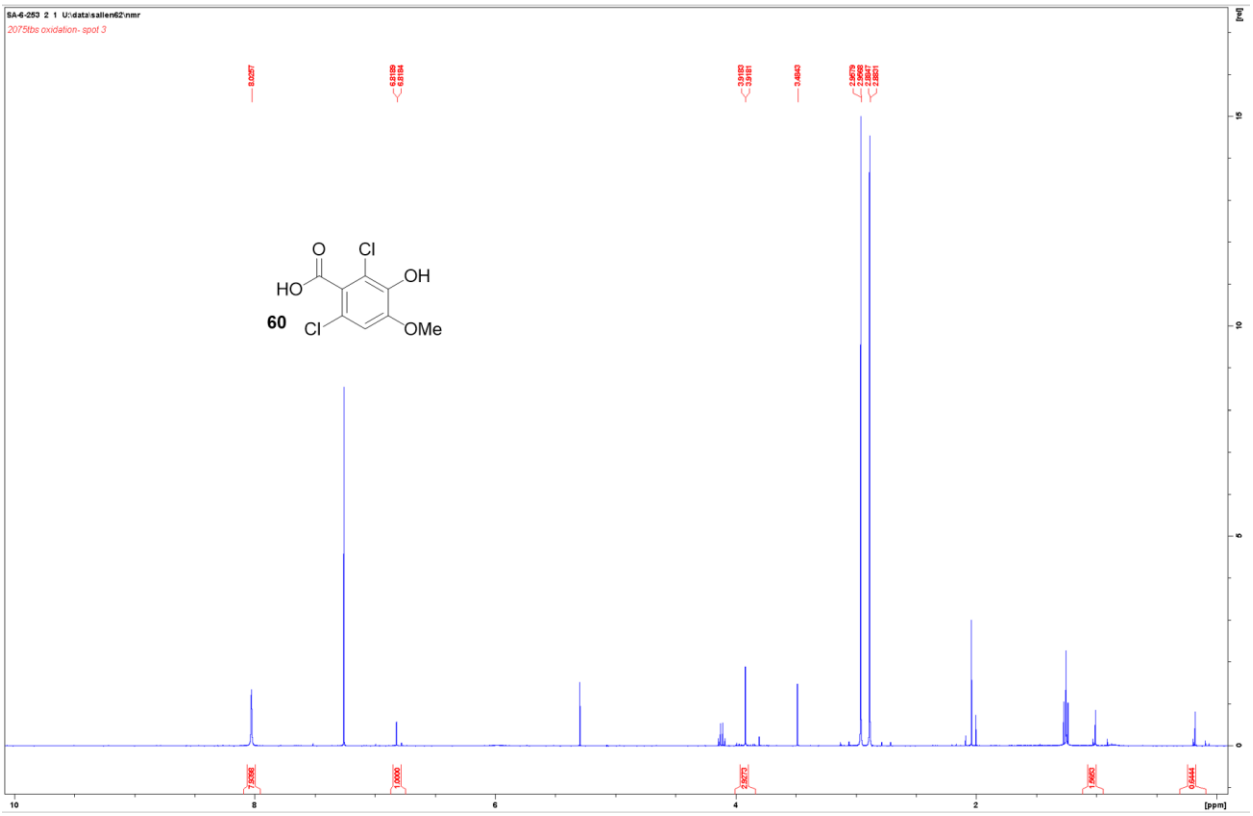
Appendix Figure 74. ^{1}H NMR Spectrum of **57.**



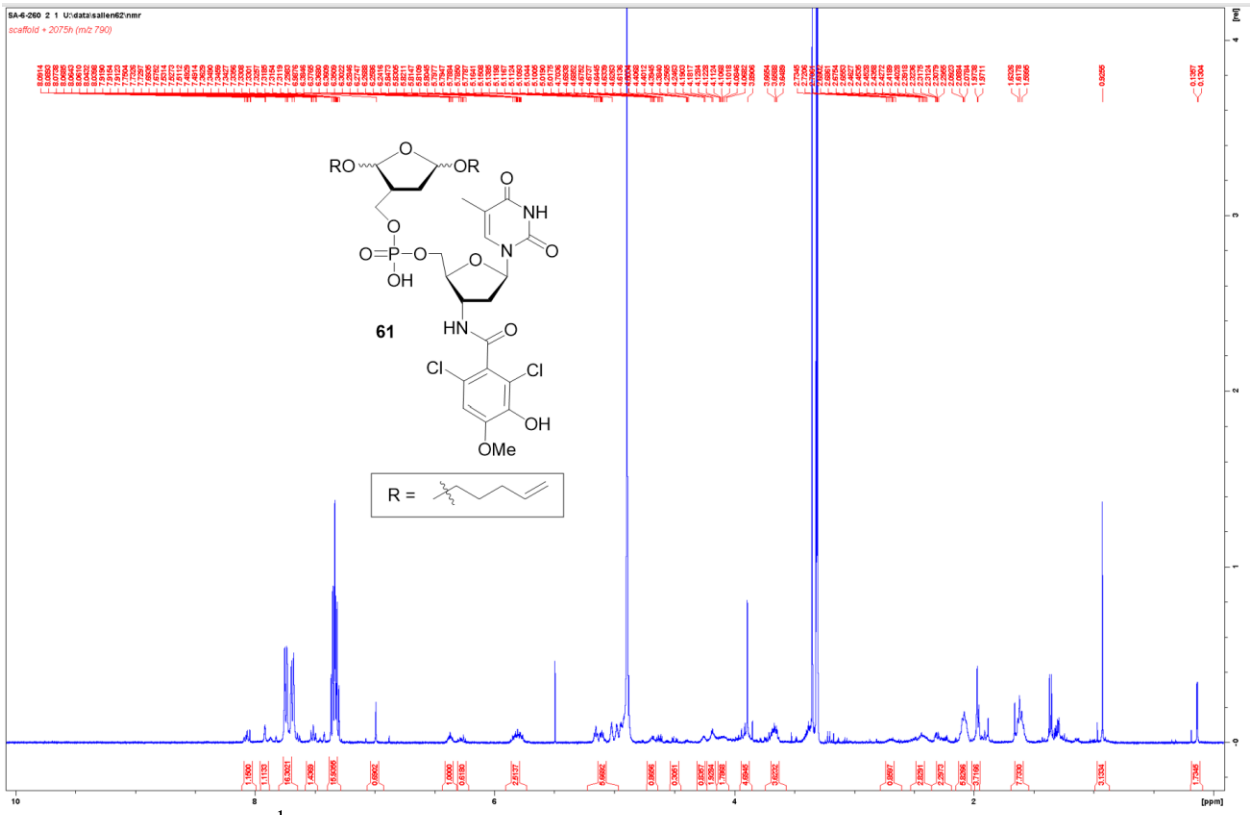
Appendix Figure 75. ^{31}P NMR Spectrum of **57.**



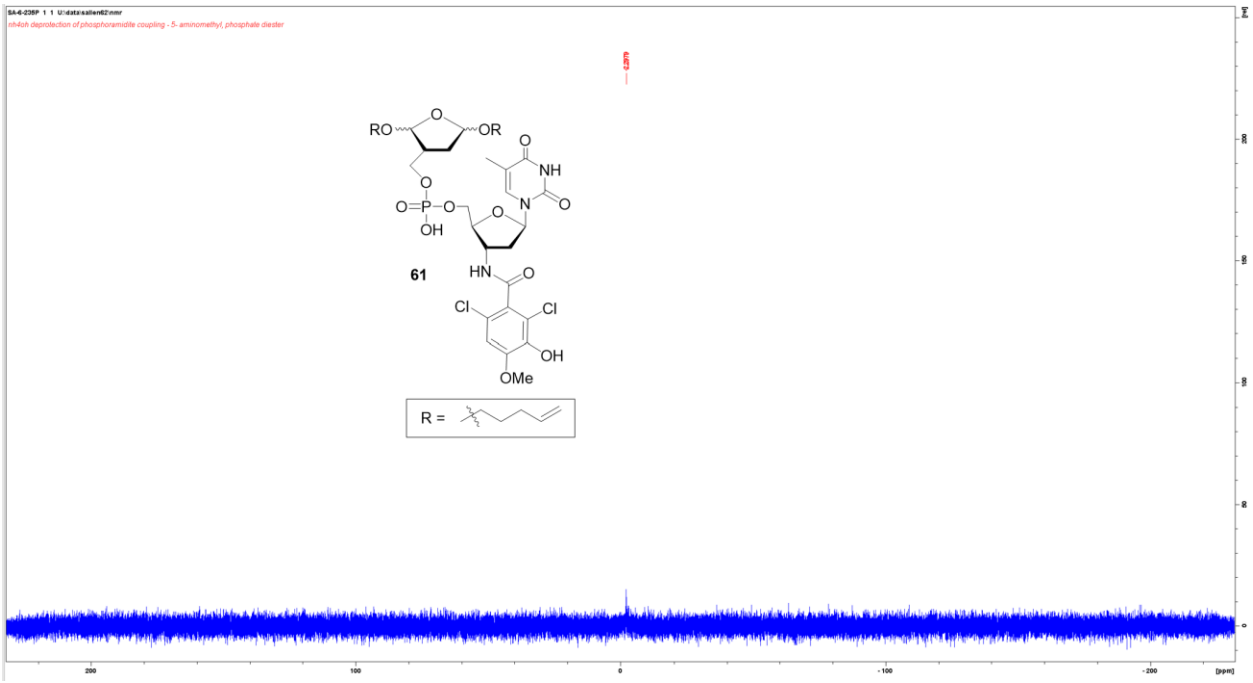
Appendix Figure 76. ^{1}H NMR Spectrum of **59.**



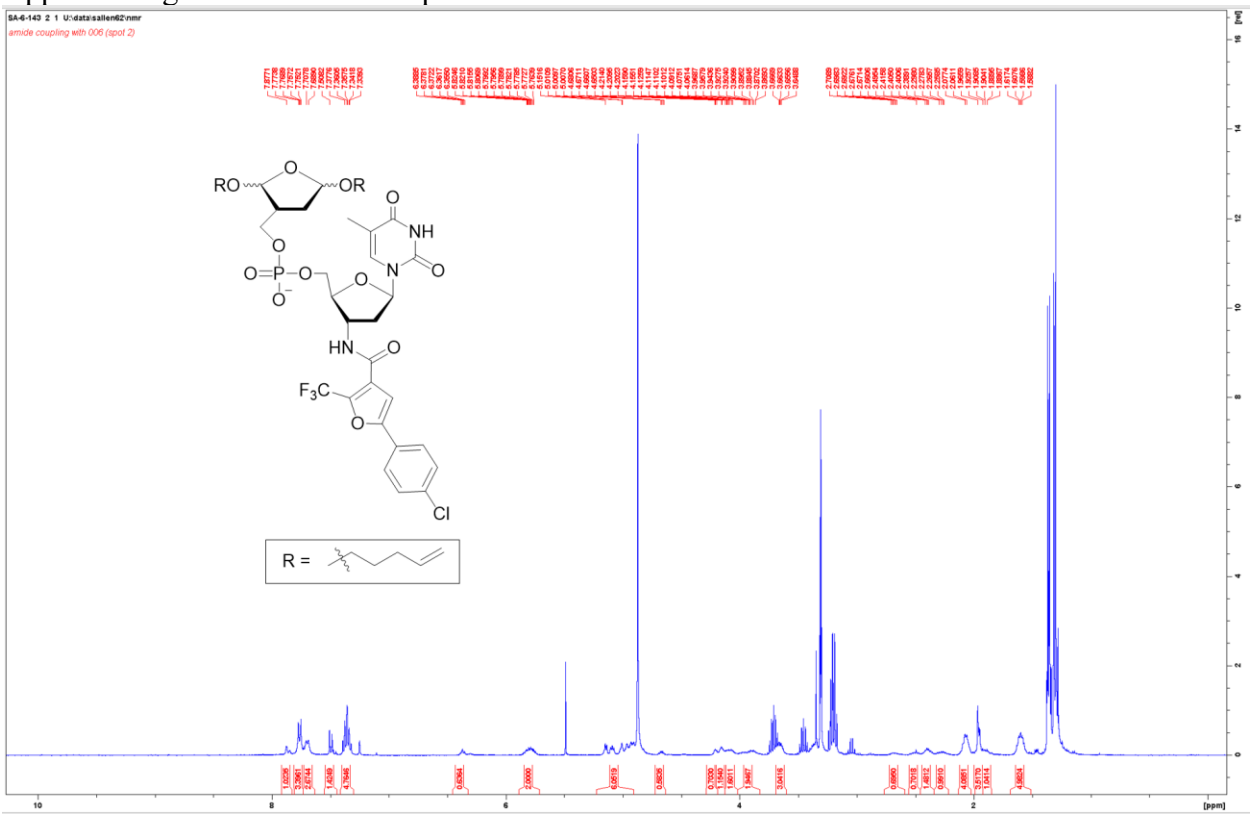
Appendix Figure 77. ^{1}H NMR Spectrum of **60.**



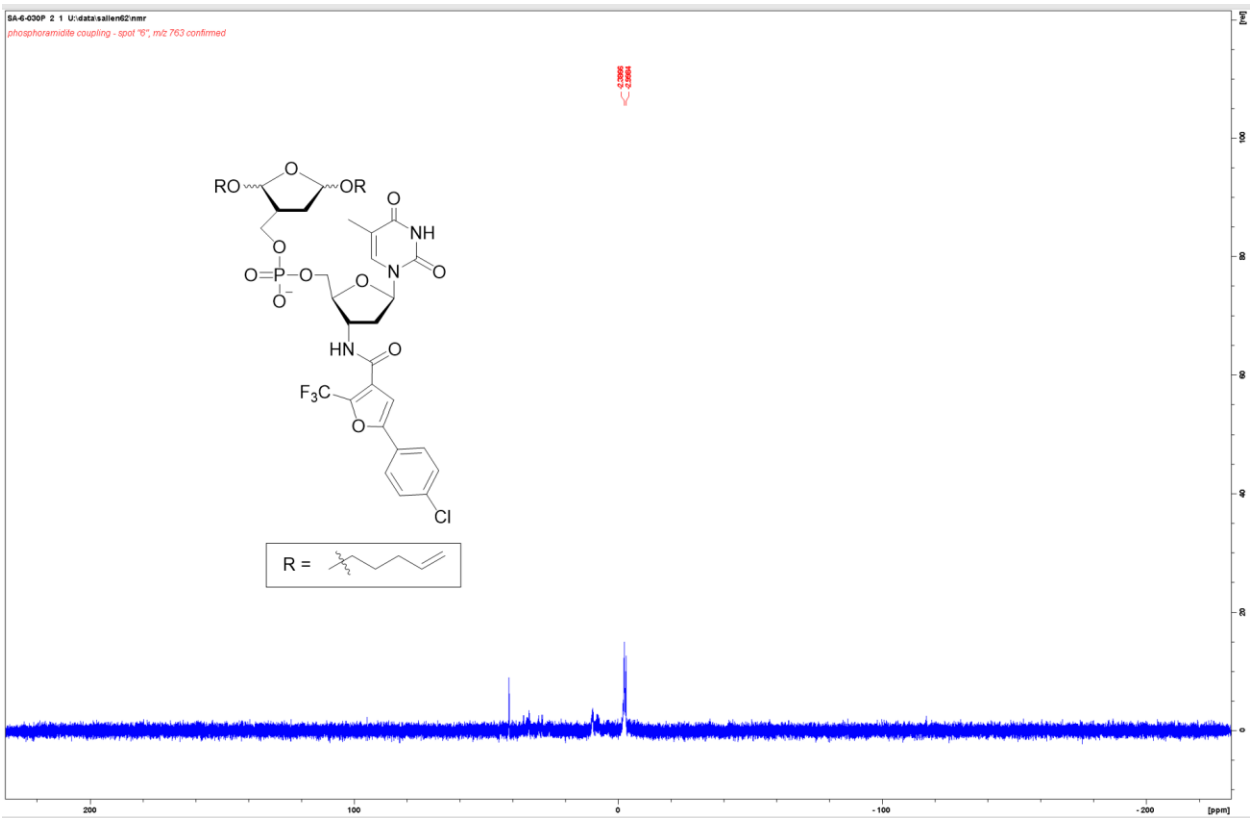
Appendix Figure 78. ^{1}H NMR Spectrum of **61.**



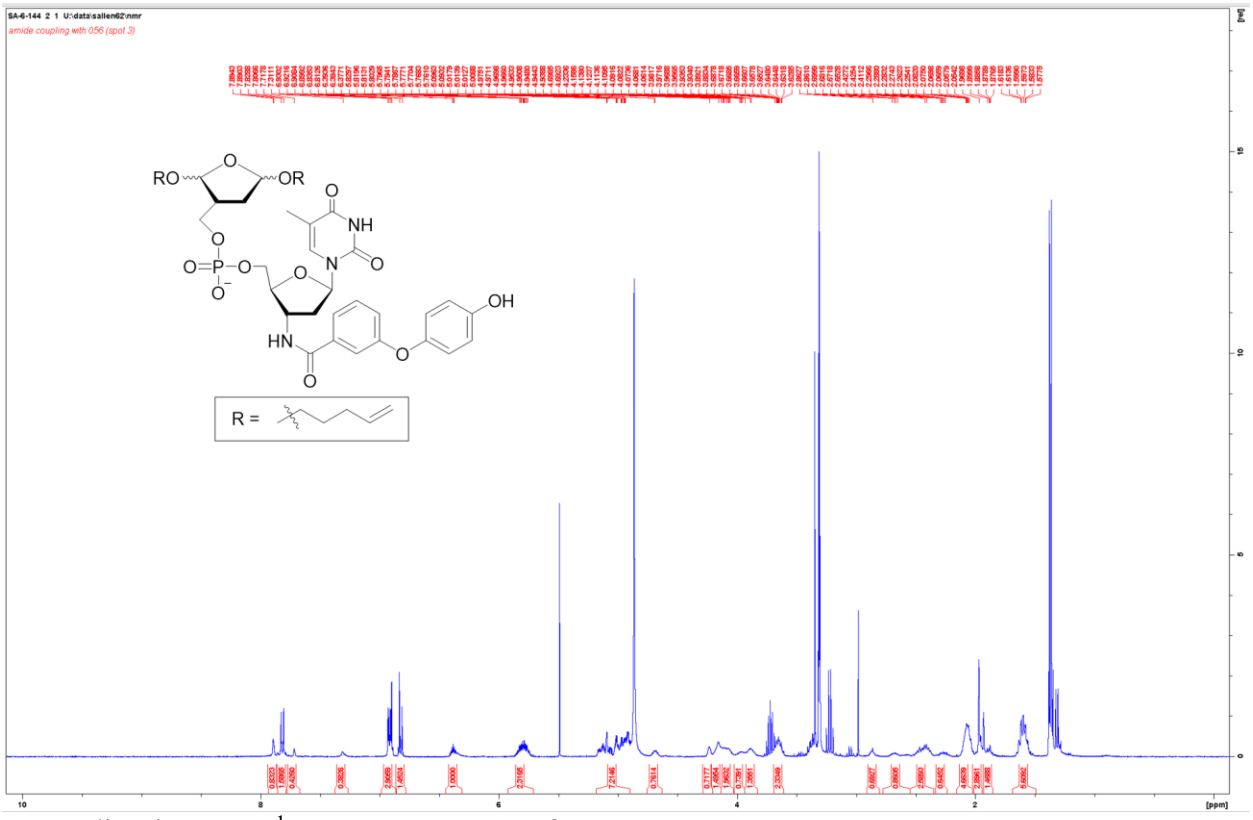
Appendix Figure 79. ^{31}P NMR Spectrum of **61.**



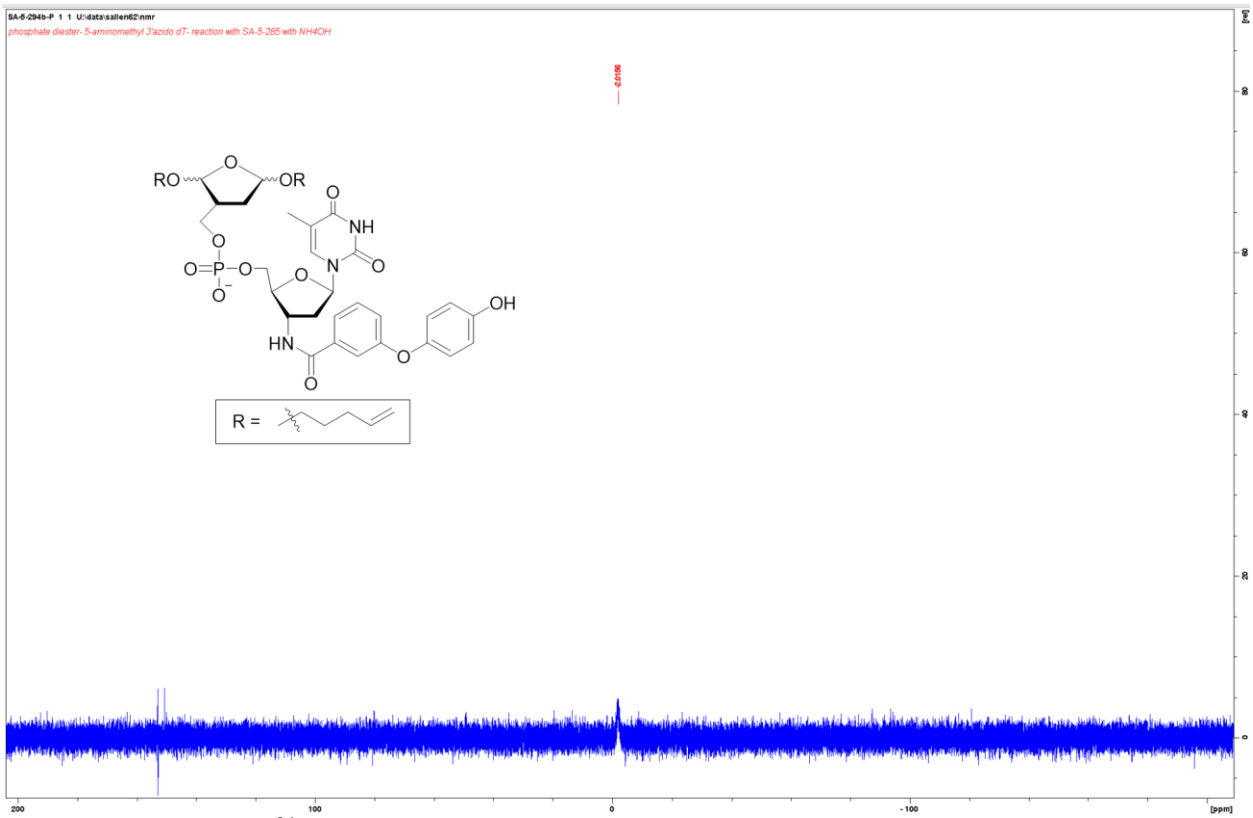
Appendix Figure 80. ^{1}H NMR Spectrum of Precursor to **63.**



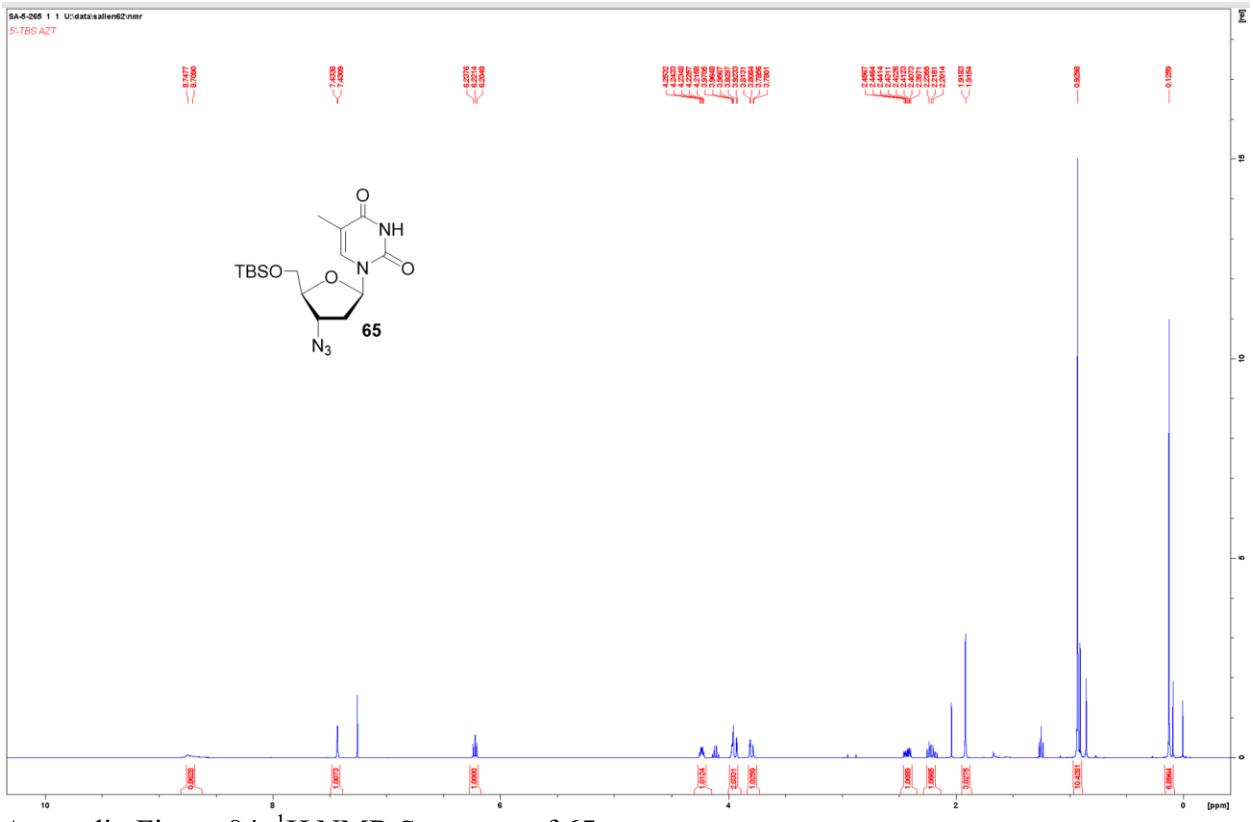
Appendix Figure 81. ^{31}P NMR Spectrum of Precursor to **63.**



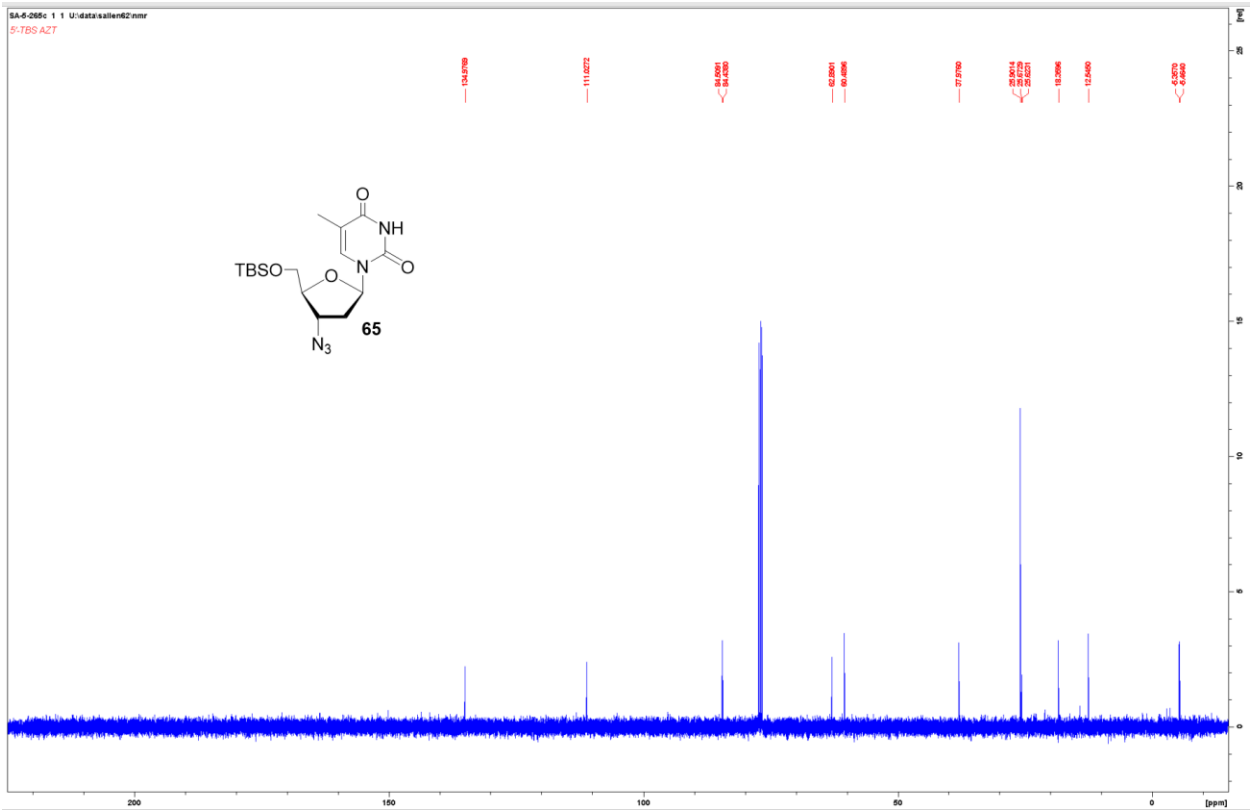
Appendix Figure 82. ^{1}H NMR Spectrum of Precursor to **64.**



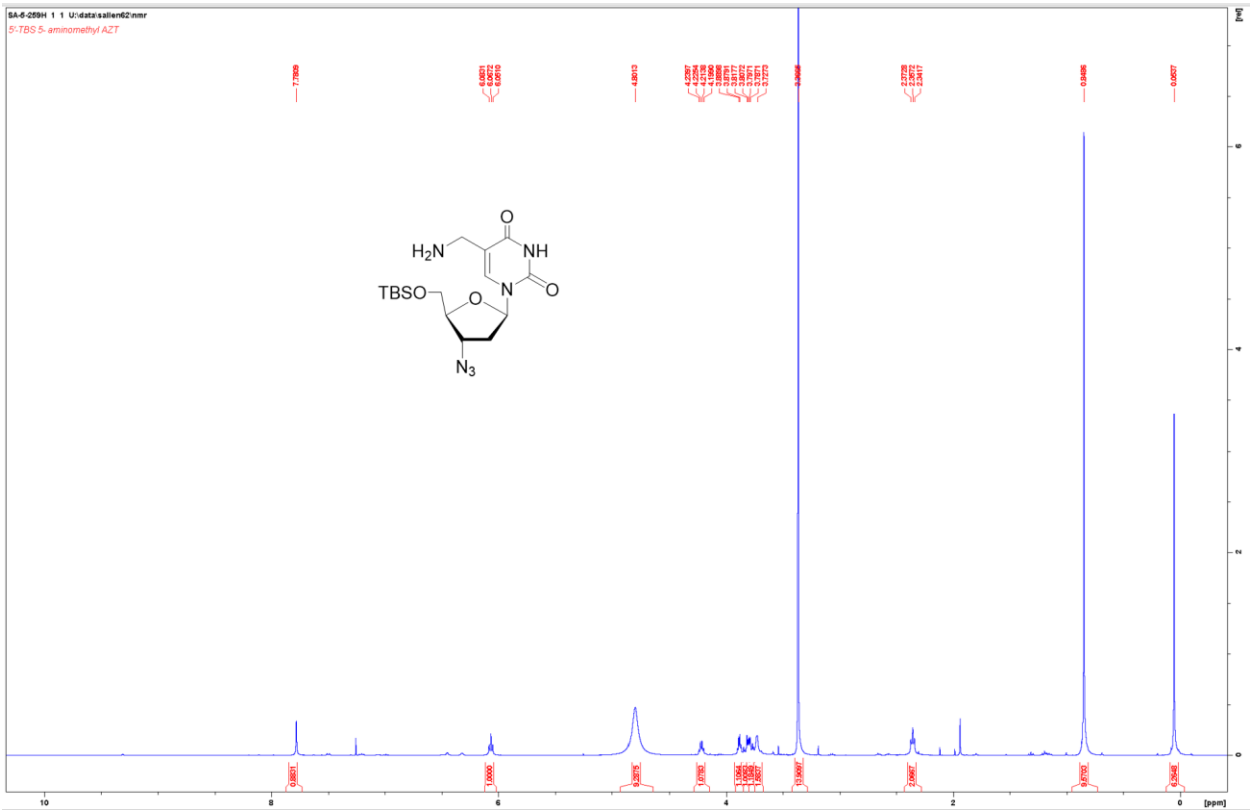
Appendix Figure 83. ^{31}P NMR Spectrum of Precursor to **64.**



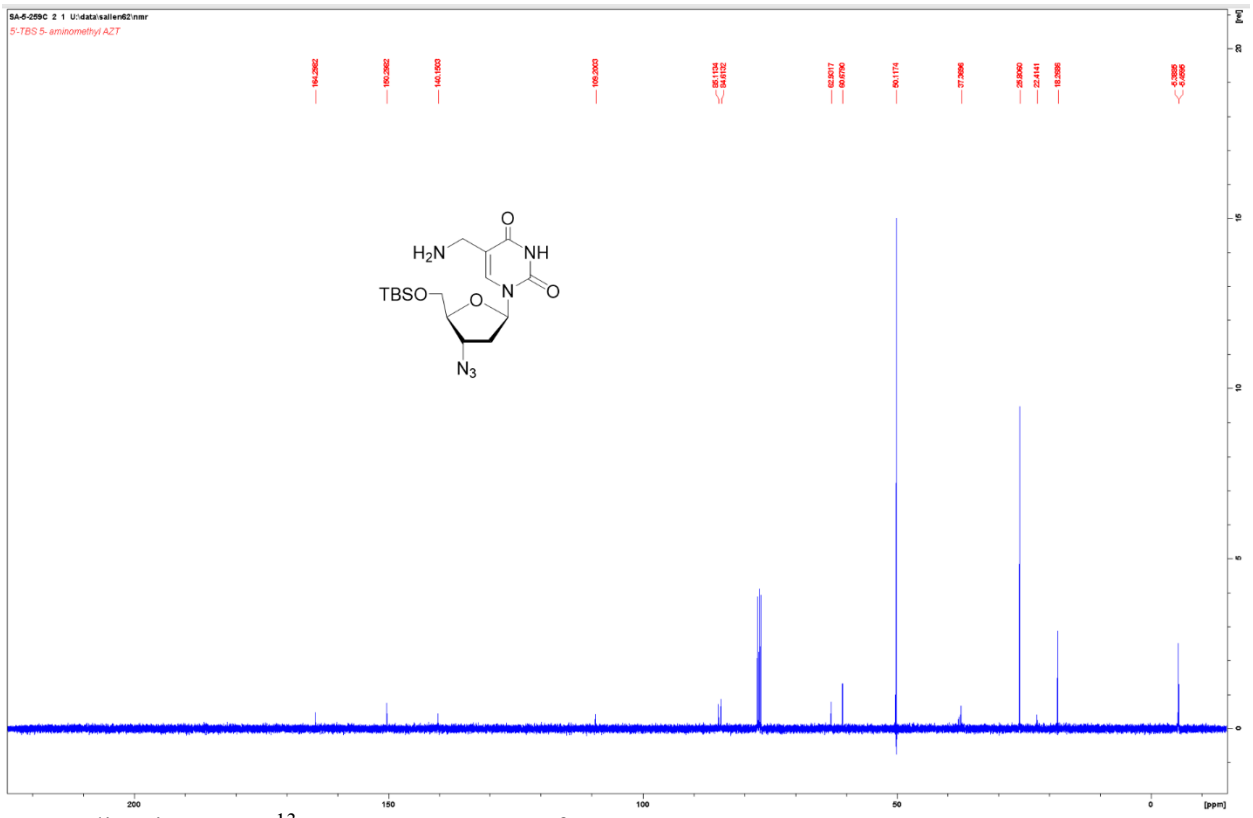
Appendix Figure 84. ^{1}H NMR Spectrum of **65.**



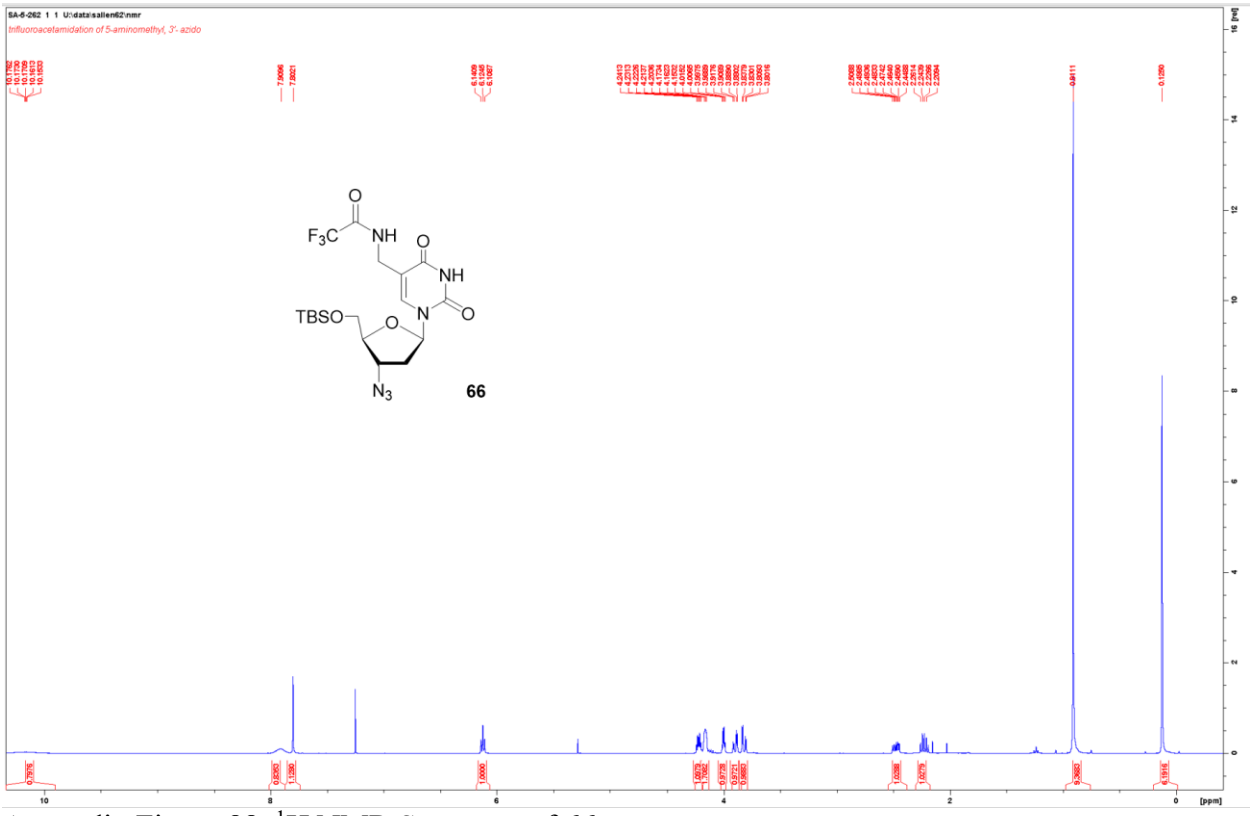
Appendix Figure 85. ^{13}C NMR Spectrum of **65.**



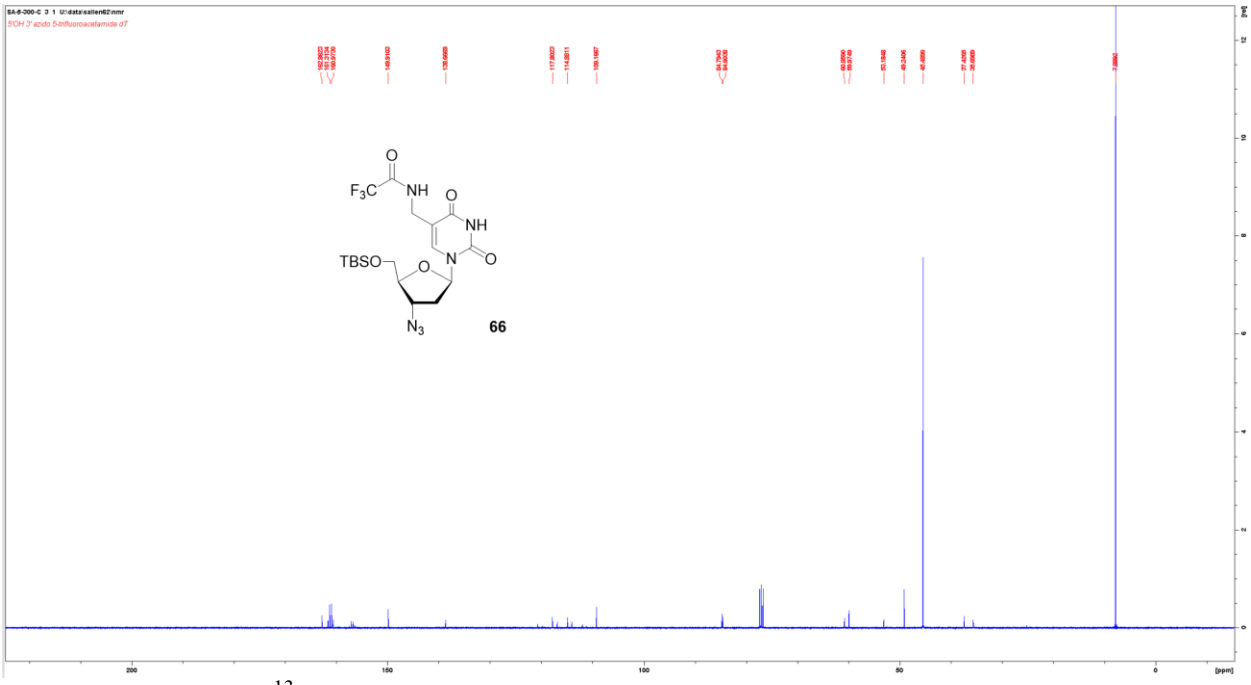
Appendix Figure 86. ^{1}H NMR Spectrum of Precursor to **66.**



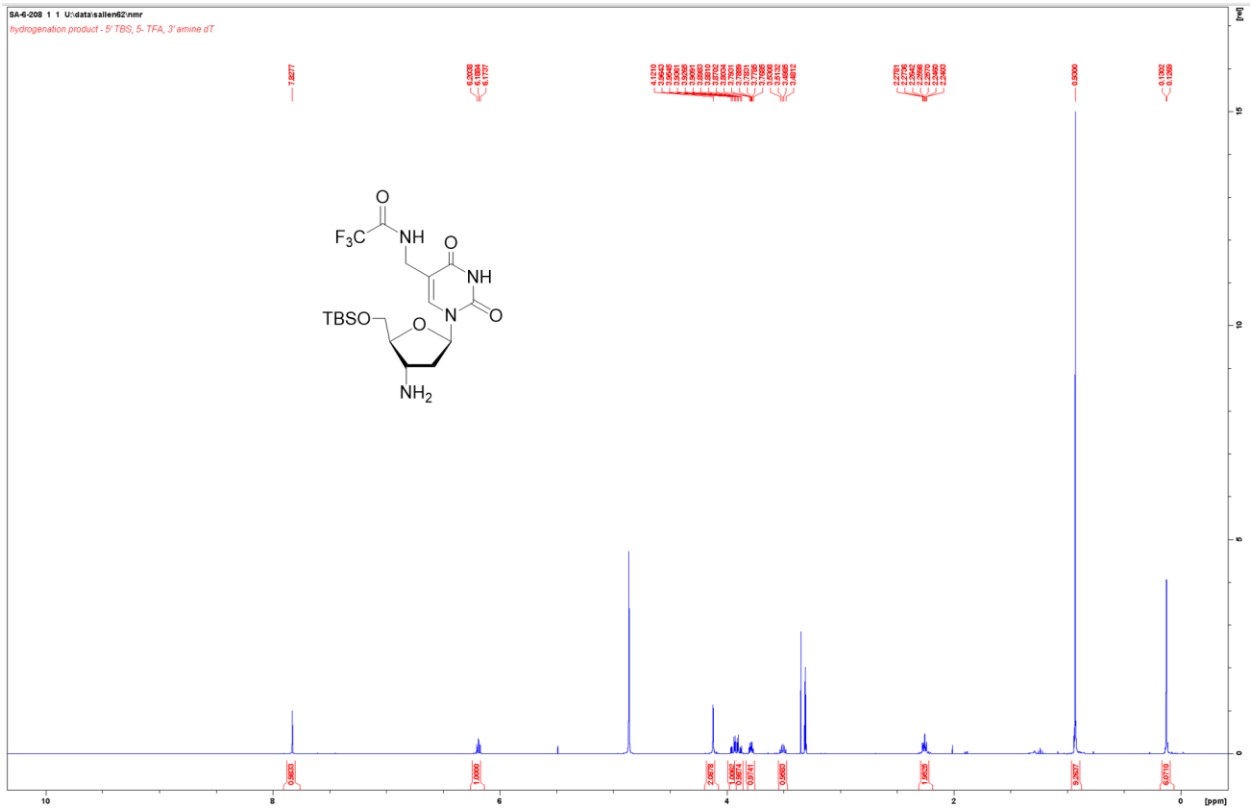
Appendix Figure 87. ^{13}C NMR Spectrum of Precursor to **66.**



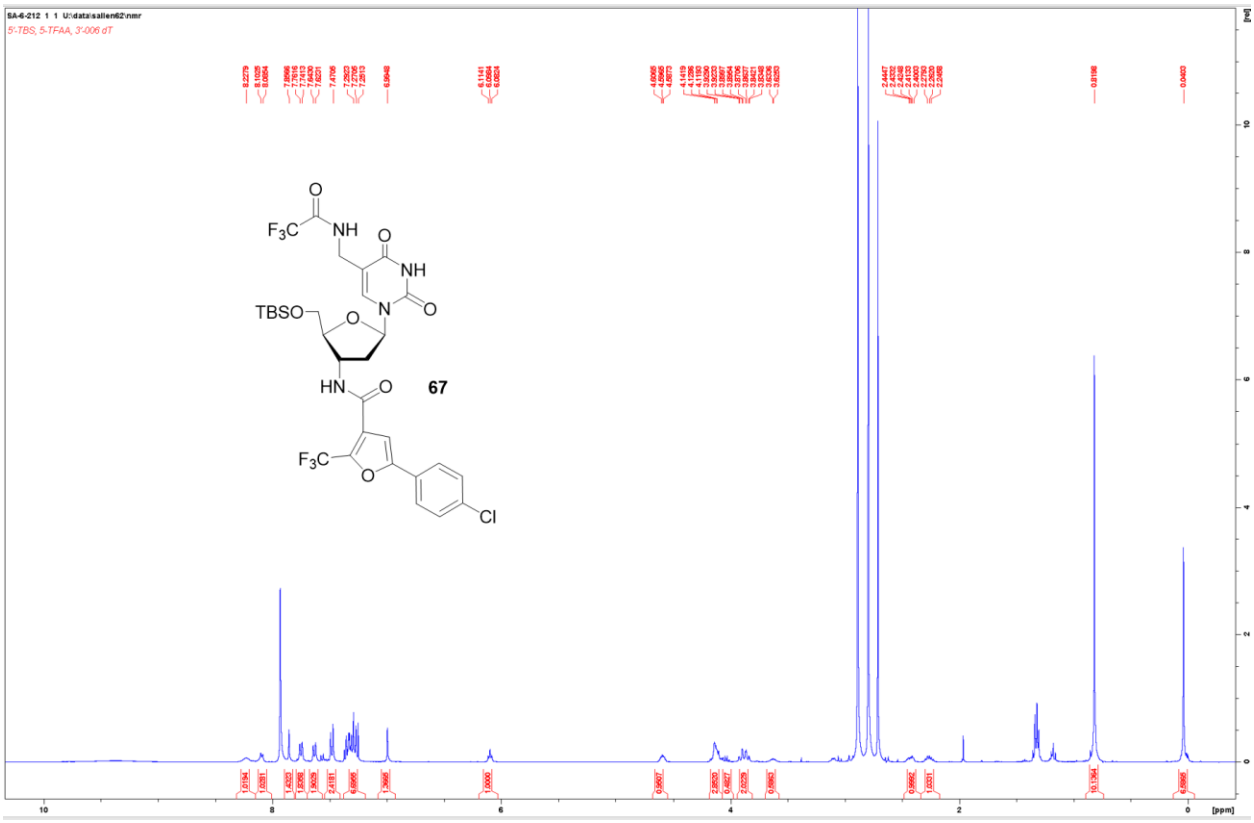
Appendix Figure 88. ^{1}H NMR Spectrum of **66.**



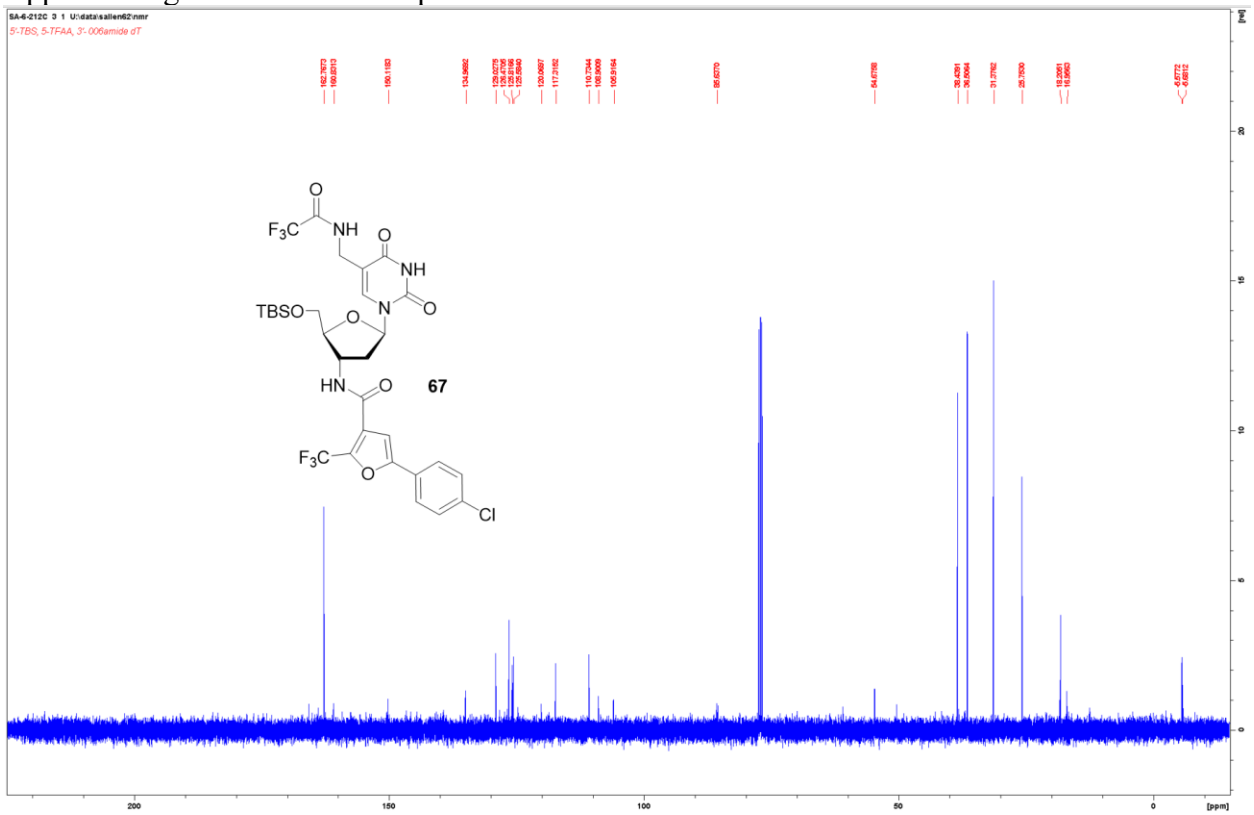
Appendix Figure 89. ^{13}C NMR Spectrum of **66.**



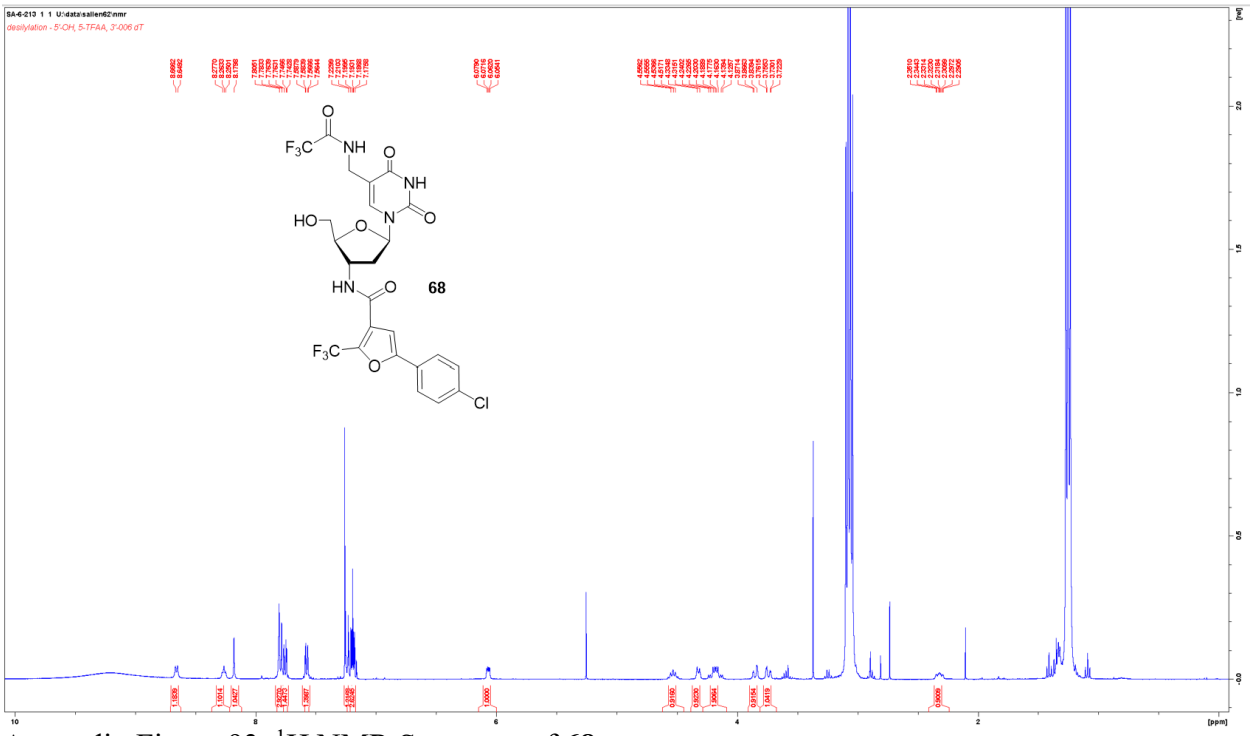
Appendix Figure 90. ^{1}H NMR Spectrum of Precursor to **67.**



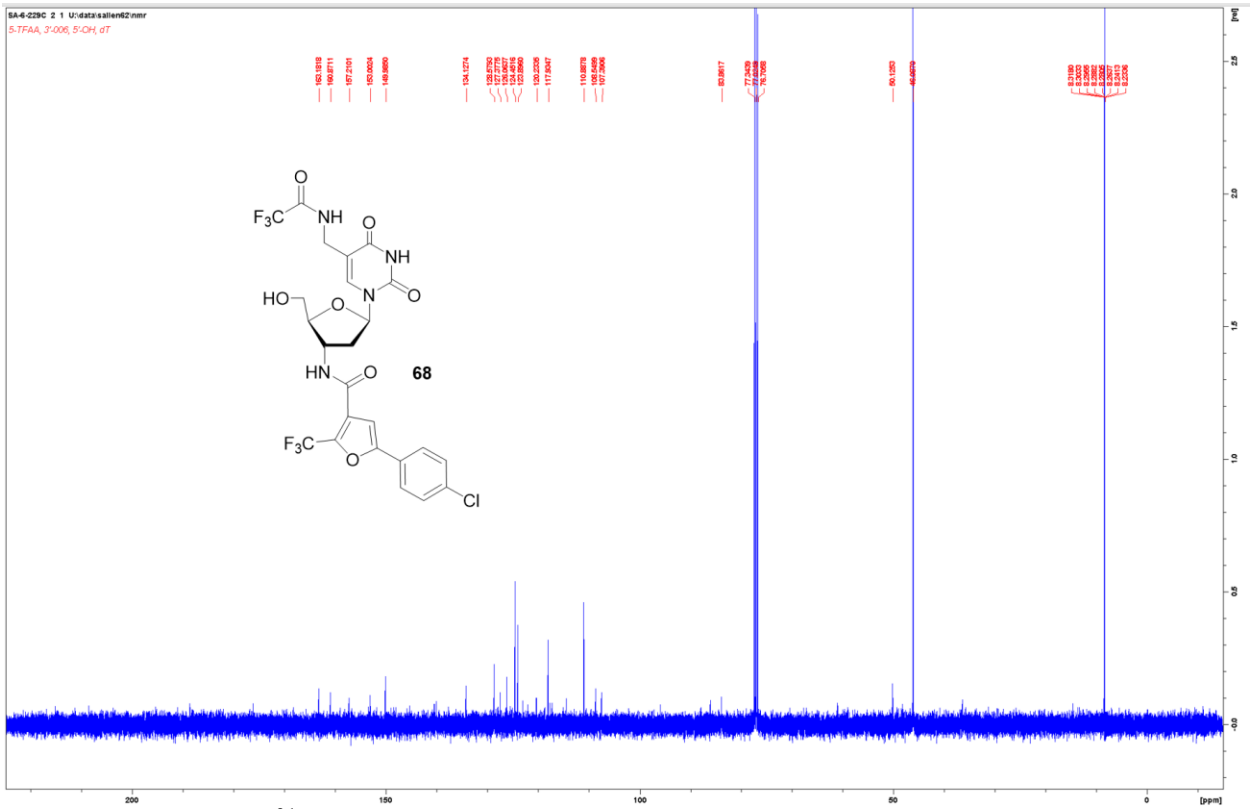
Appendix Figure 91. ^{1}H NMR Spectrum of **67.**



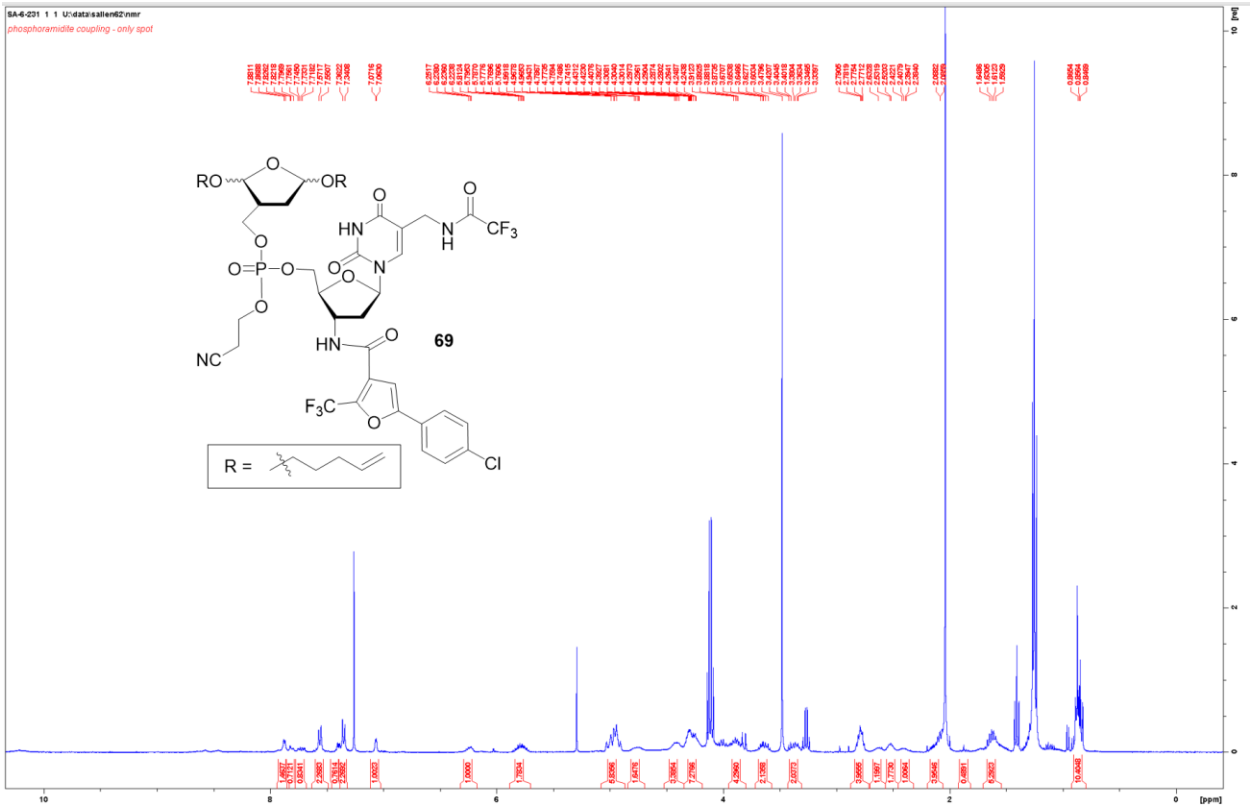
Appendix Figure 92. ^{13}C NMR Spectrum of **67.**



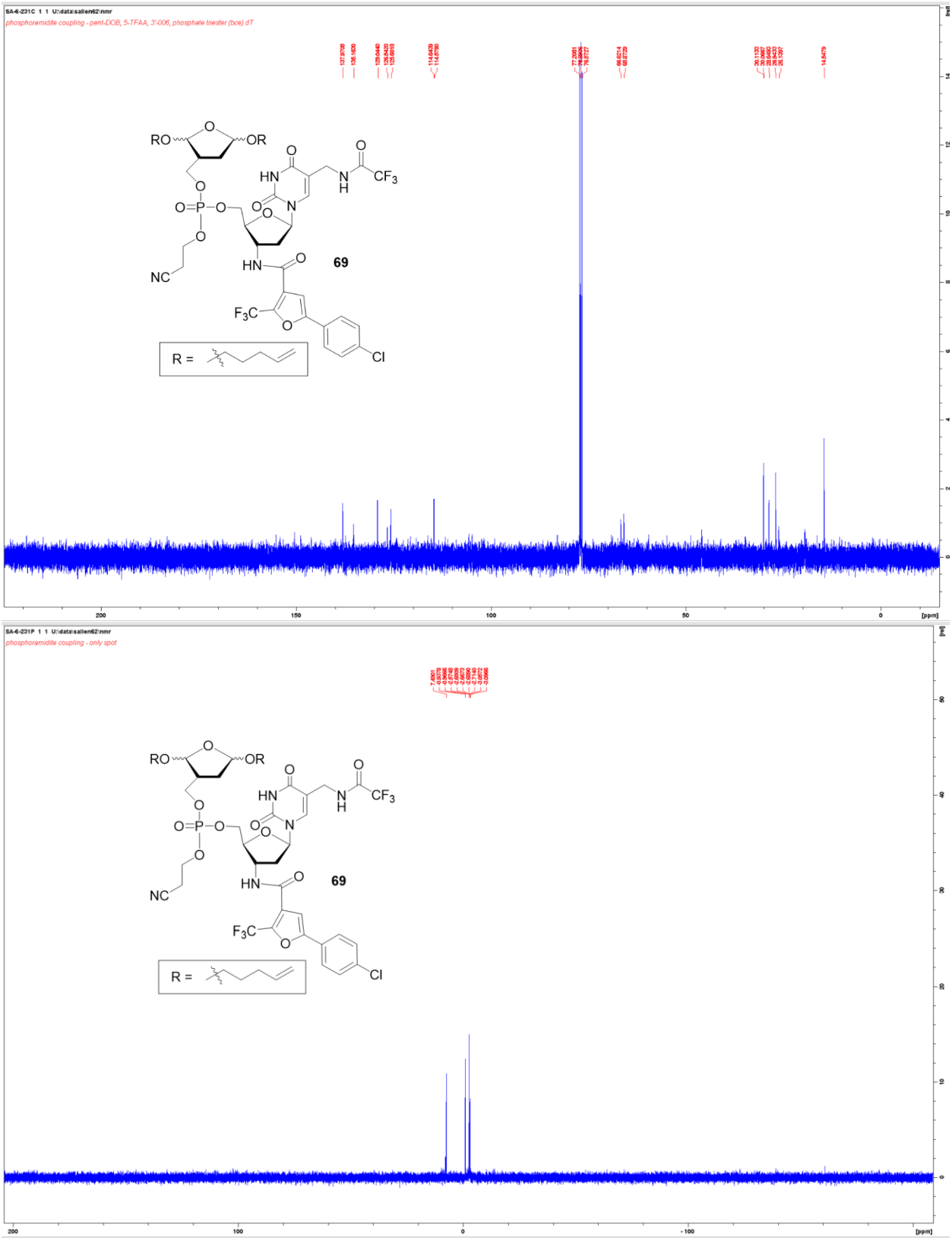
Appendix Figure 93. ^{1}H NMR Spectrum of **68.**



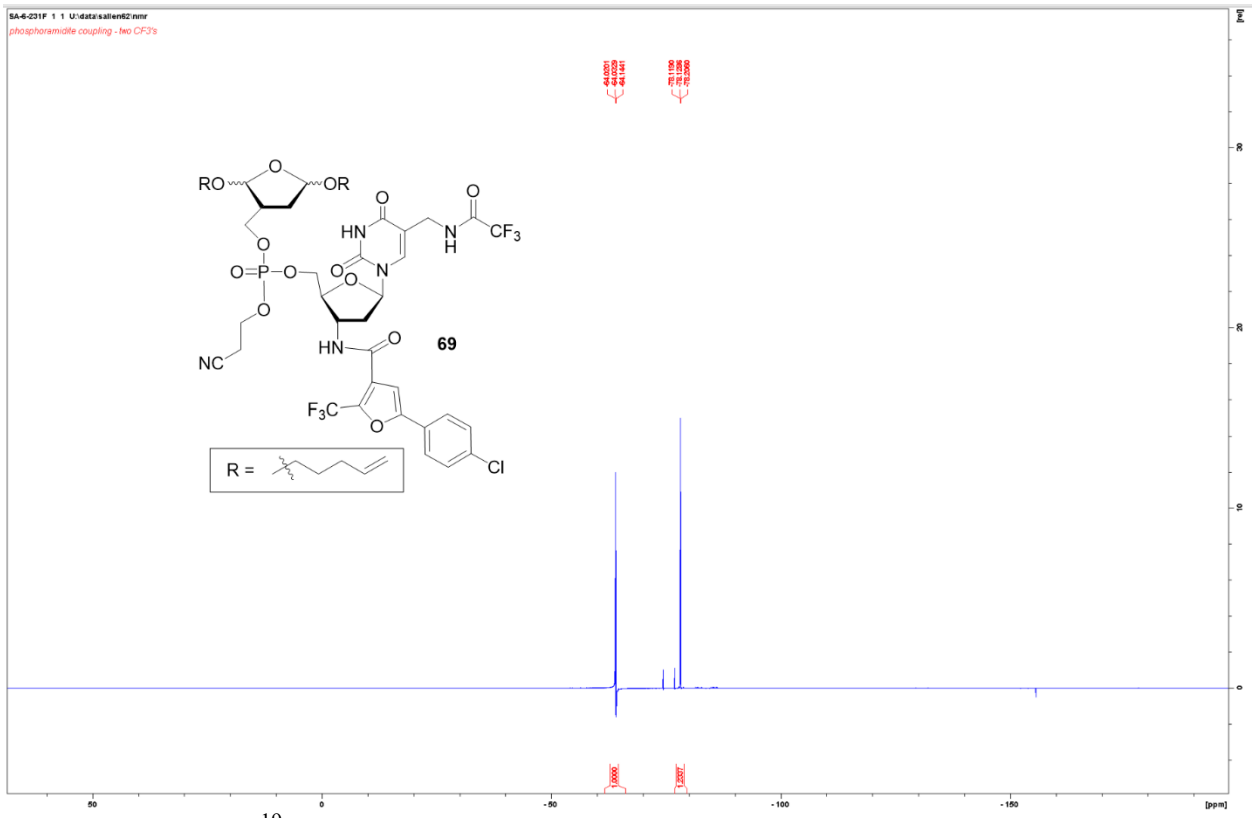
Appendix Figure 94. ^{31}P NMR Spectrum of **68.**



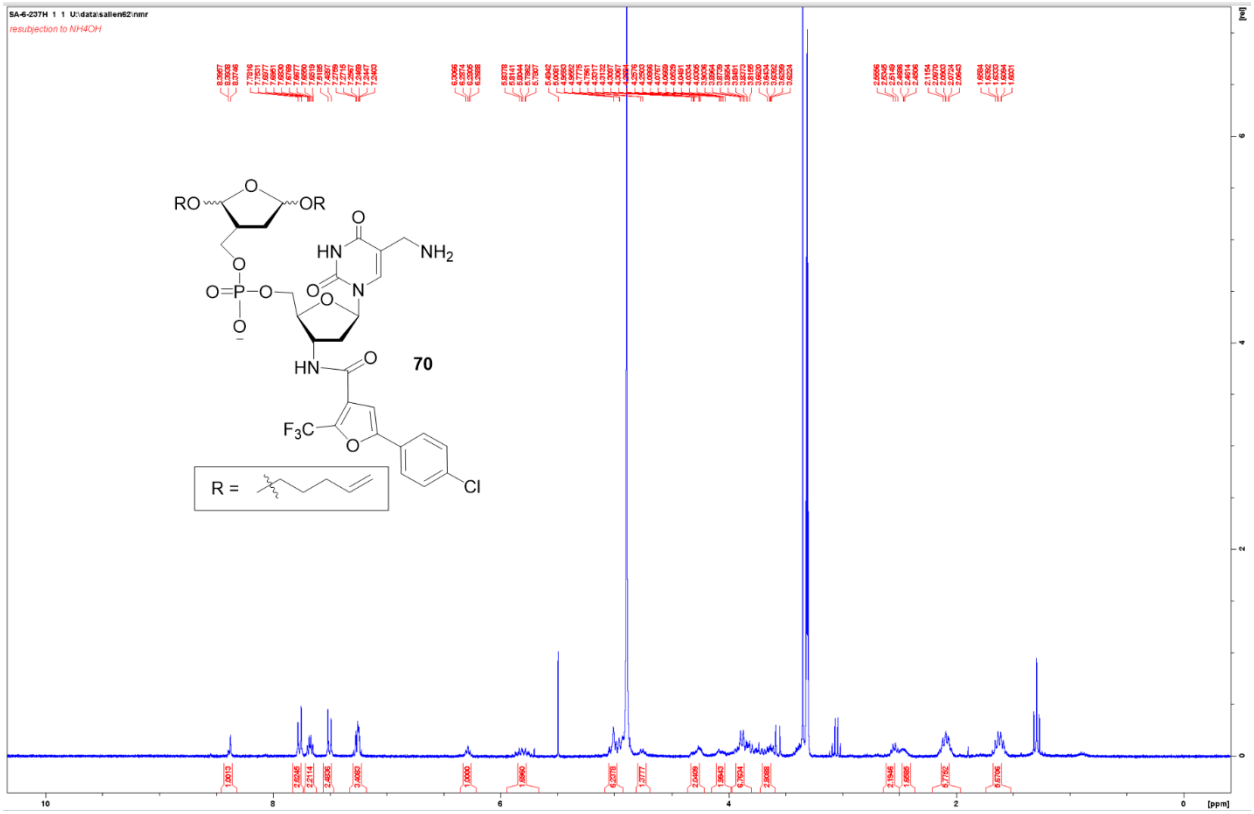
Appendix Figure 95. ^{1}H NMR Spectrum of **69.**



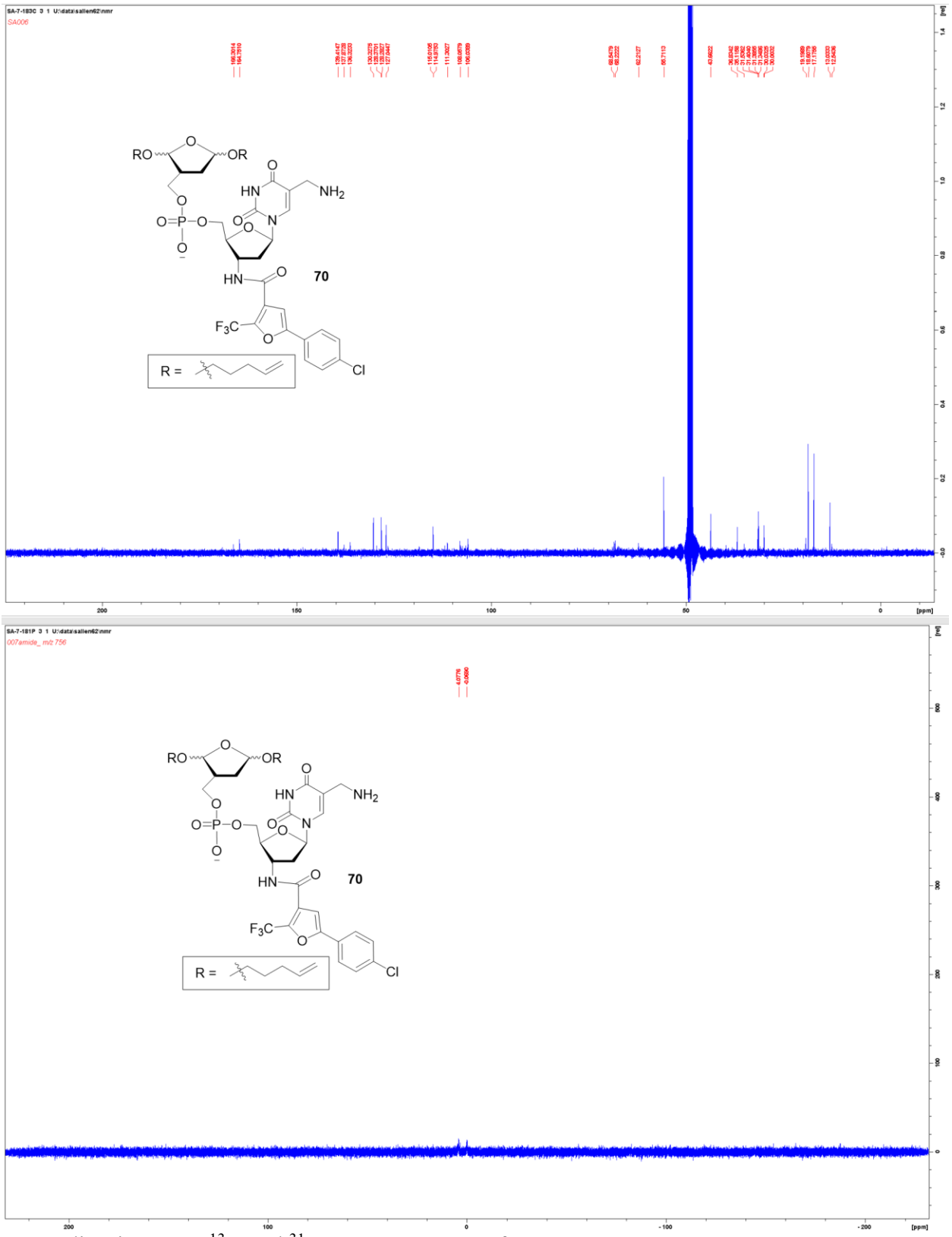
Appendix Figure 96. ^{13}C and ^{31}P NMR Spectrum of **69.**



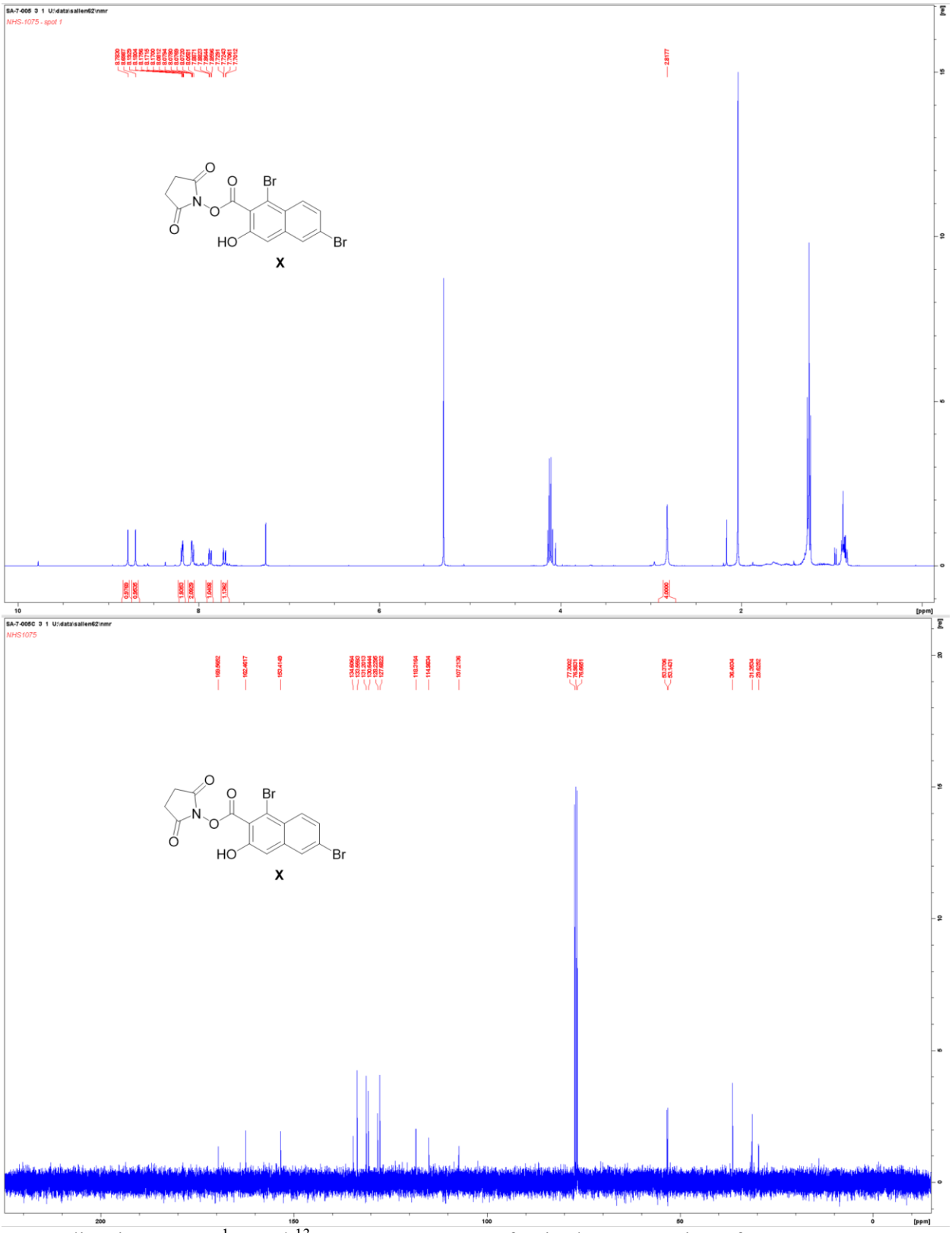
Appendix Figure 97. ^{19}F NMR Spectrum of **69.**



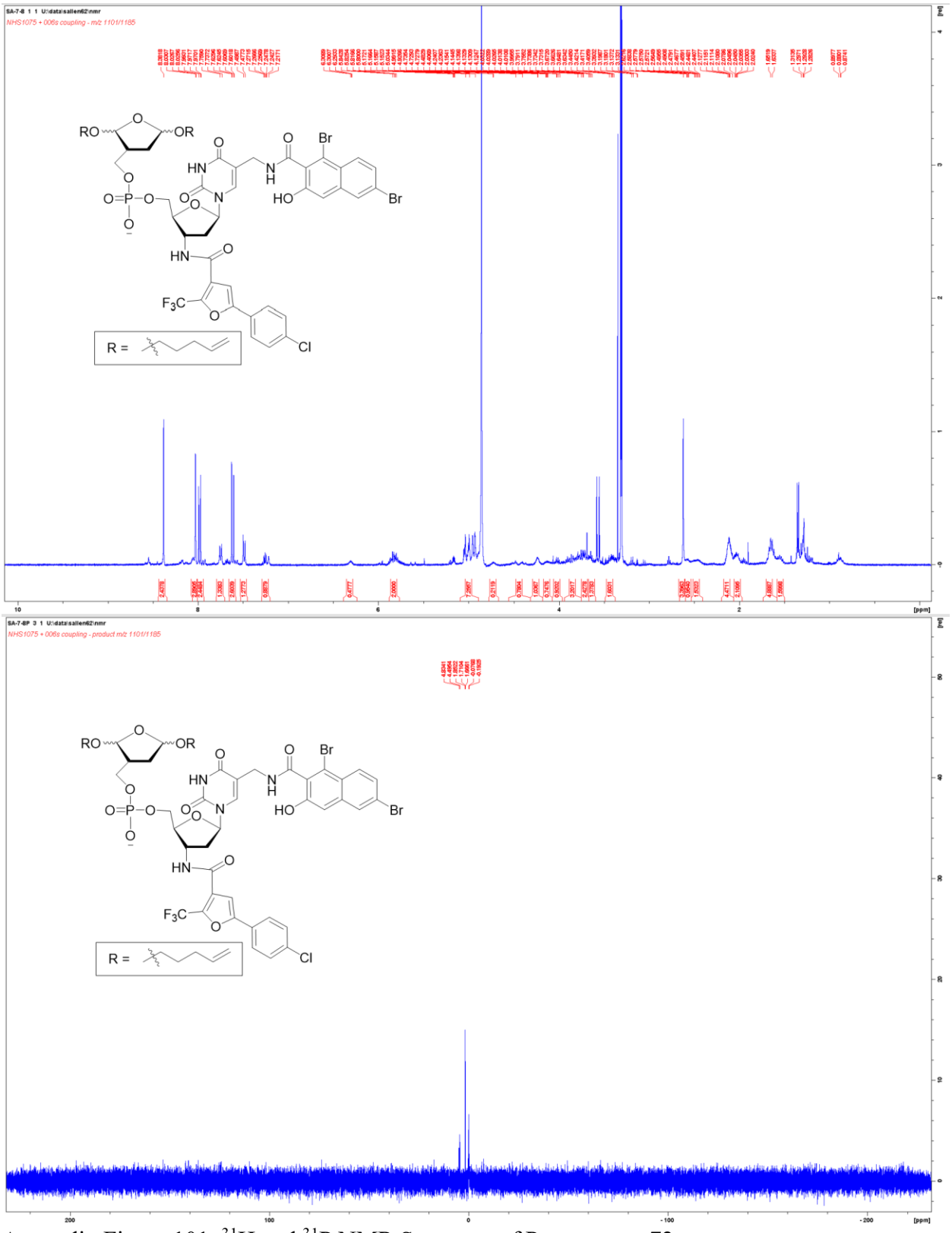
Appendix Figure 98. ^{1}H NMR Spectrum of **70**.



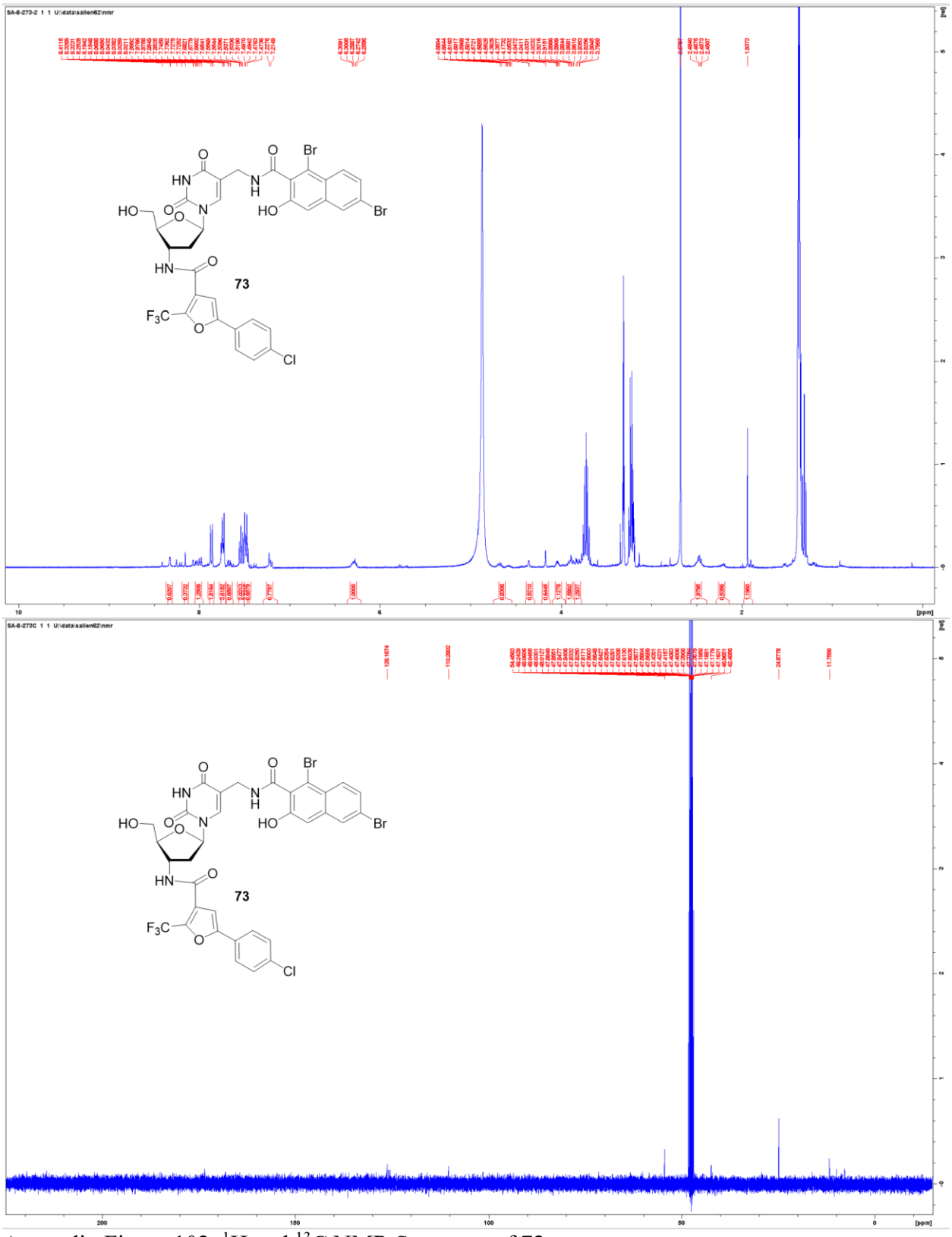
Appendix Figure 99. ^{13}C and ^{31}P NMR Spectrum of **70**.



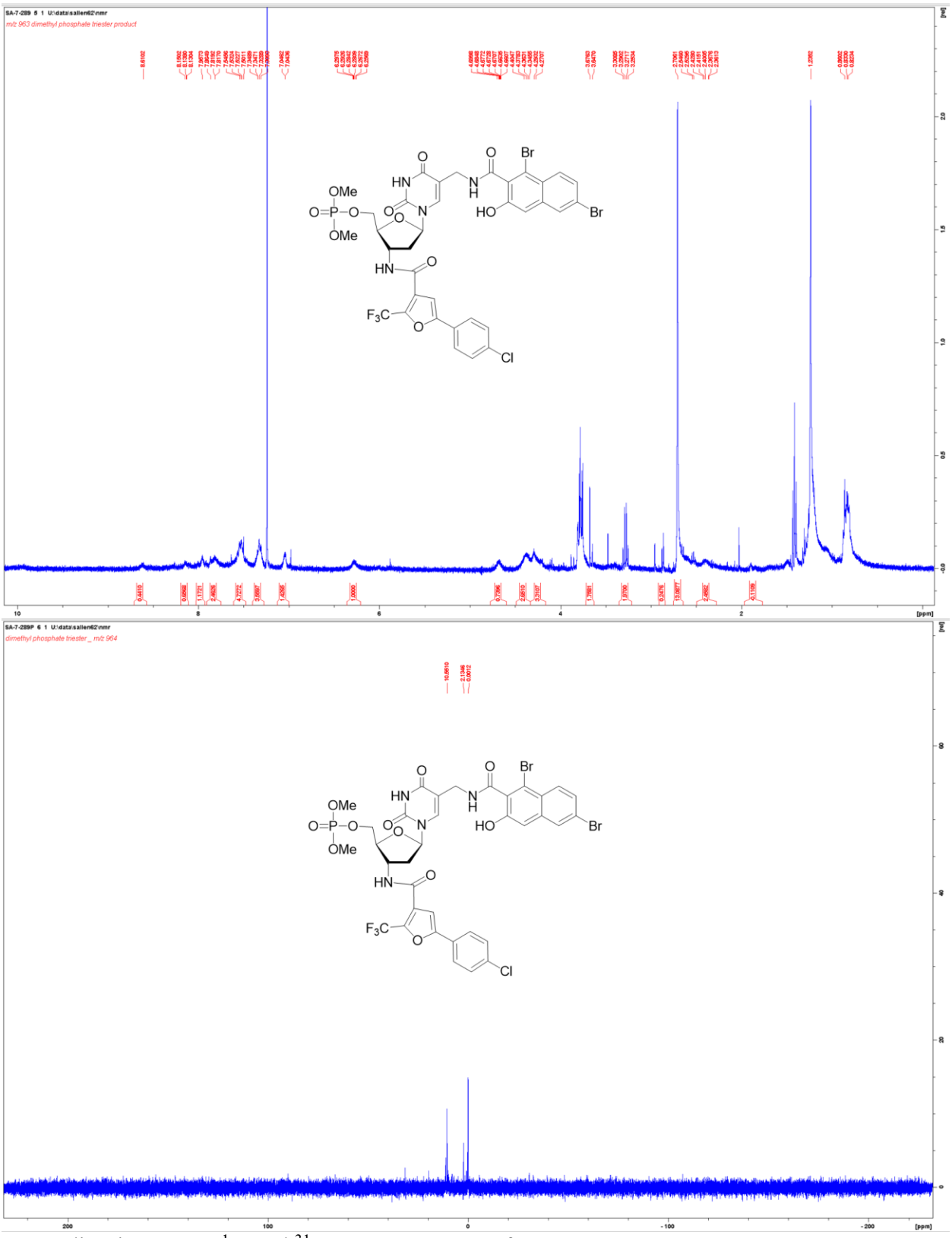
Appendix Figure 100. ^{1}H and ^{13}C NMR Spectrum of **X** in the preparation of **72**.



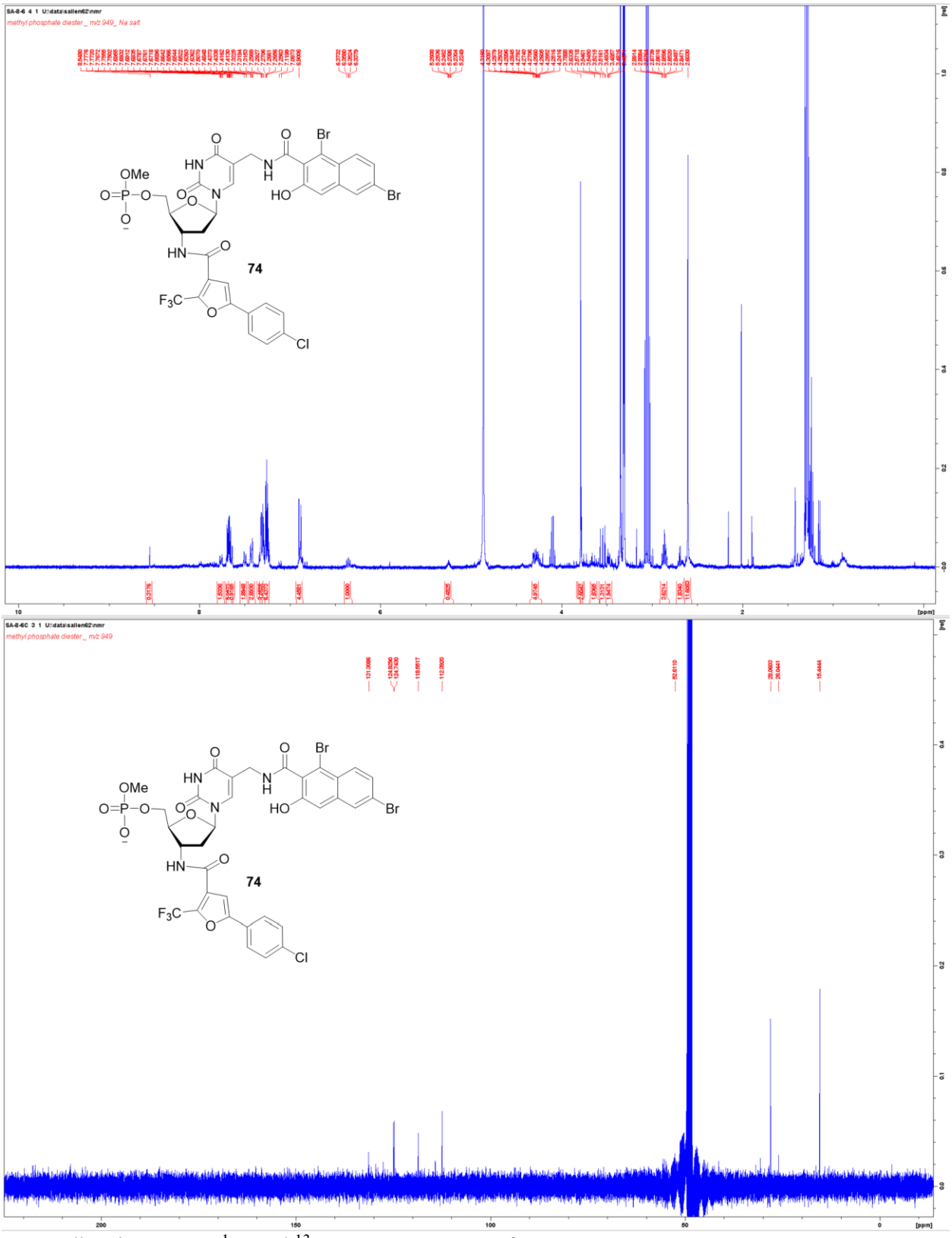
Appendix Figure 101. ^{31}H and ^{31}P NMR Spectrum of Precursor to **72**.



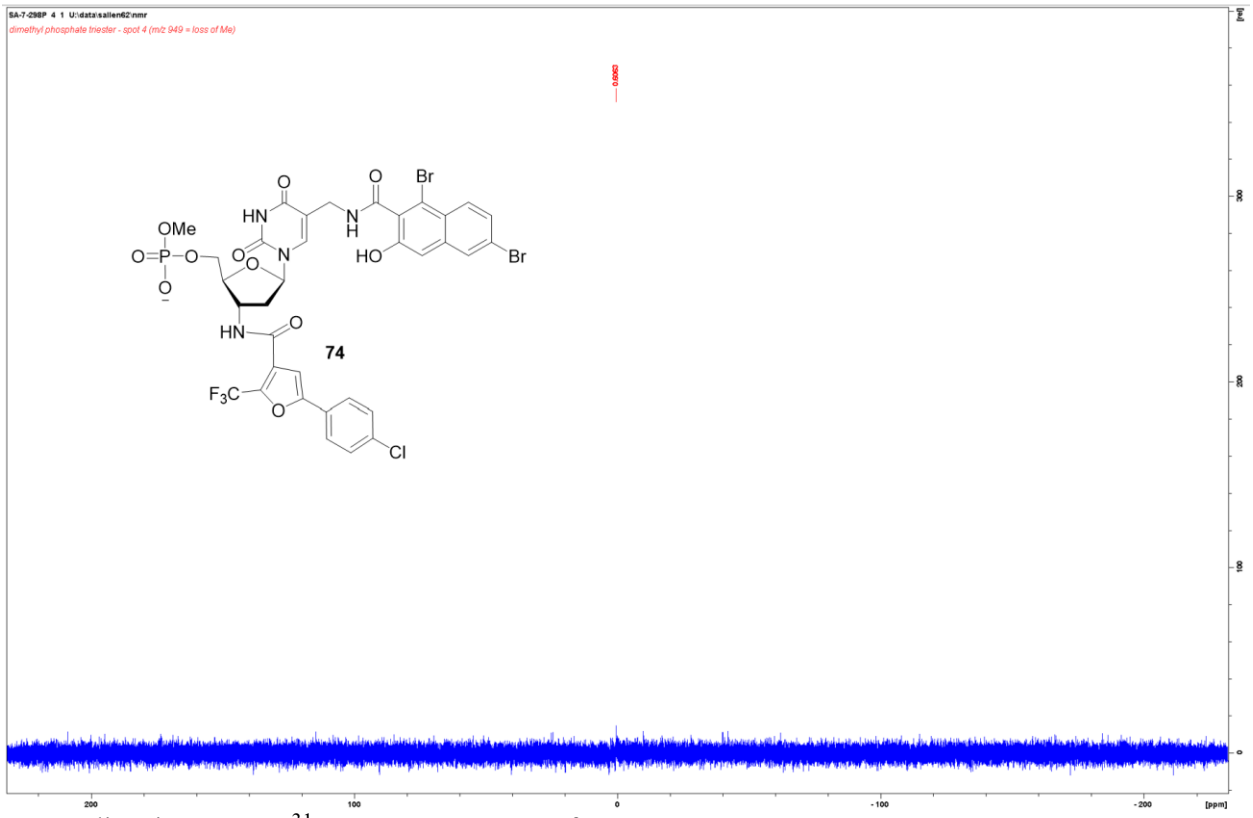
Appendix Figure 102. ^{1}H and ^{13}C NMR Spectrum of **73**.



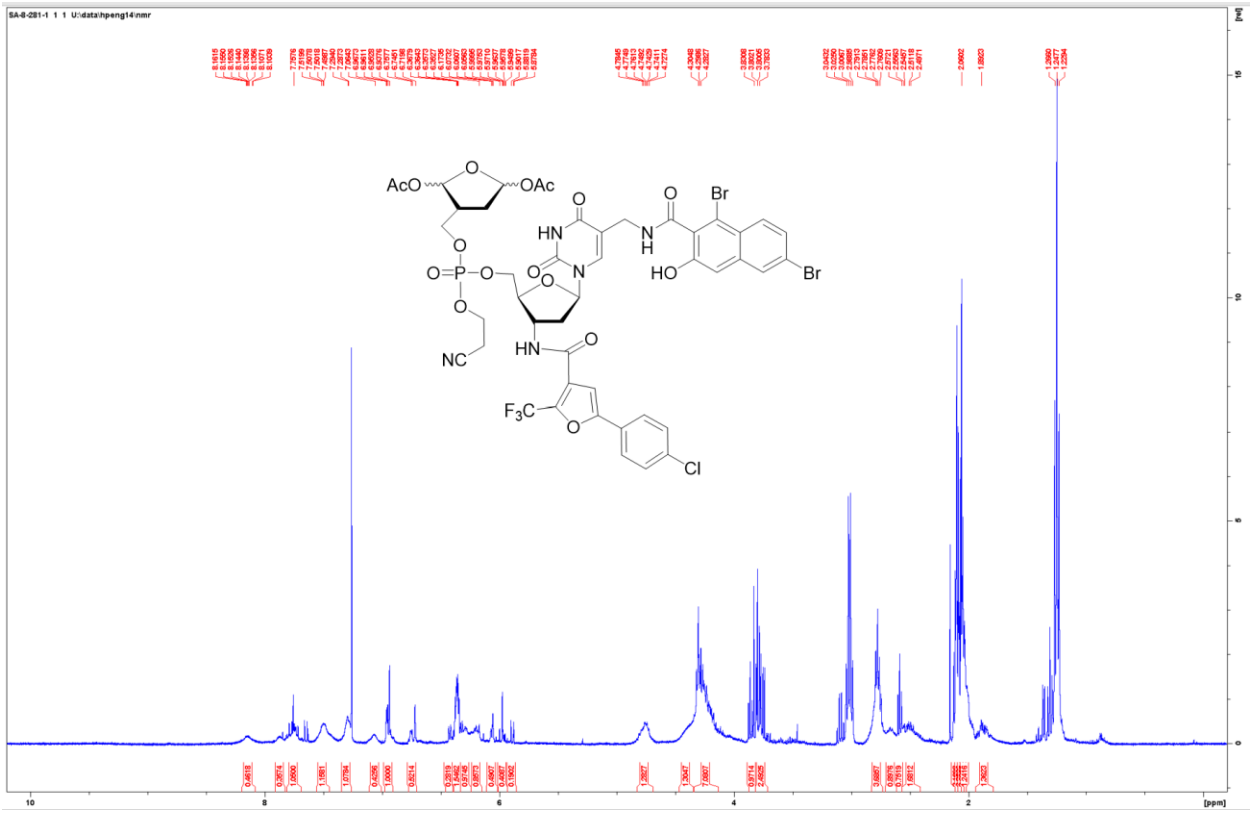
Appendix Figure 103. ^{1}H and ^{31}P NMR Spectrum of Precursor to **74**.



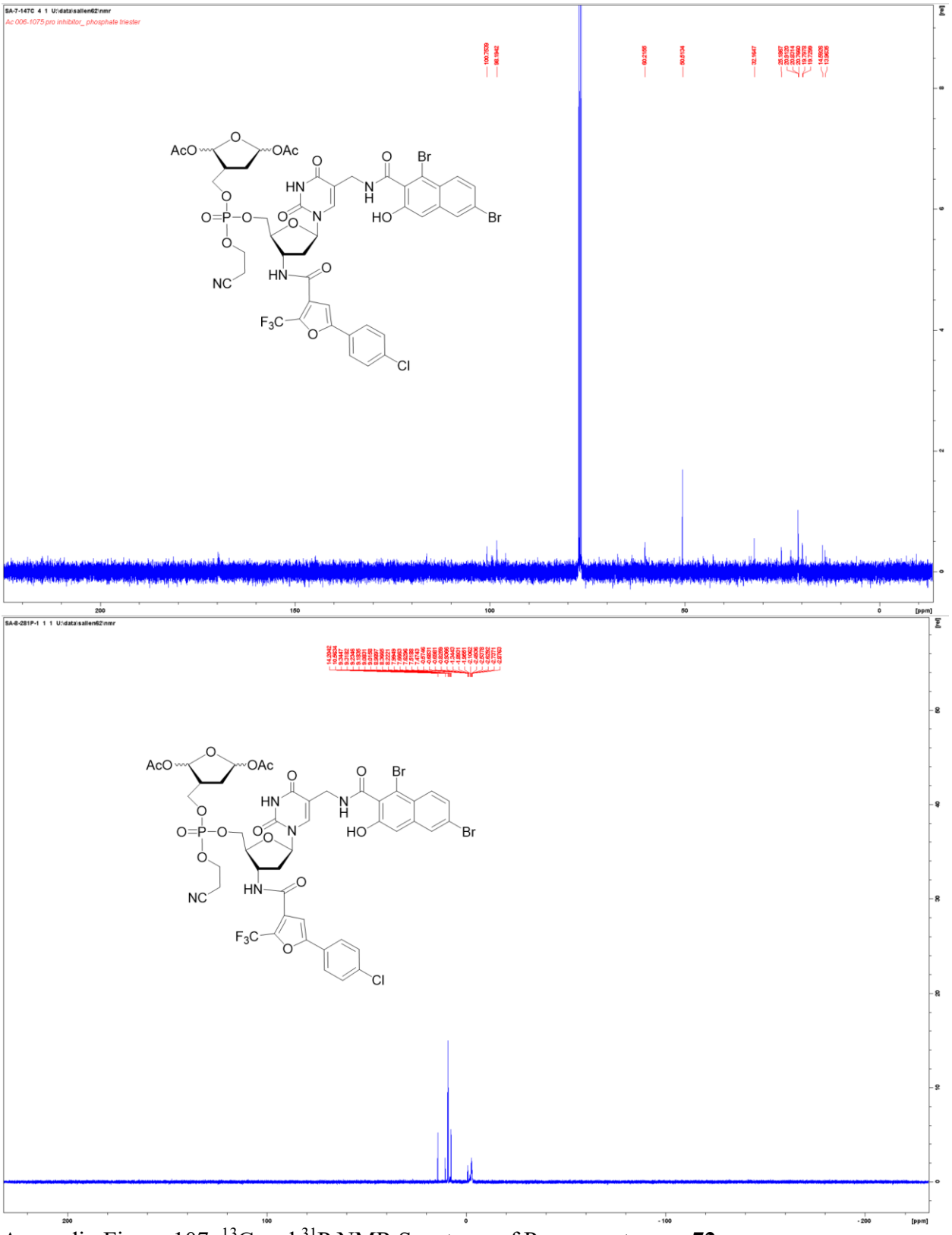
Appendix Figure 104. ^{1}H and ^{13}C NMR Spectrum of **74**.



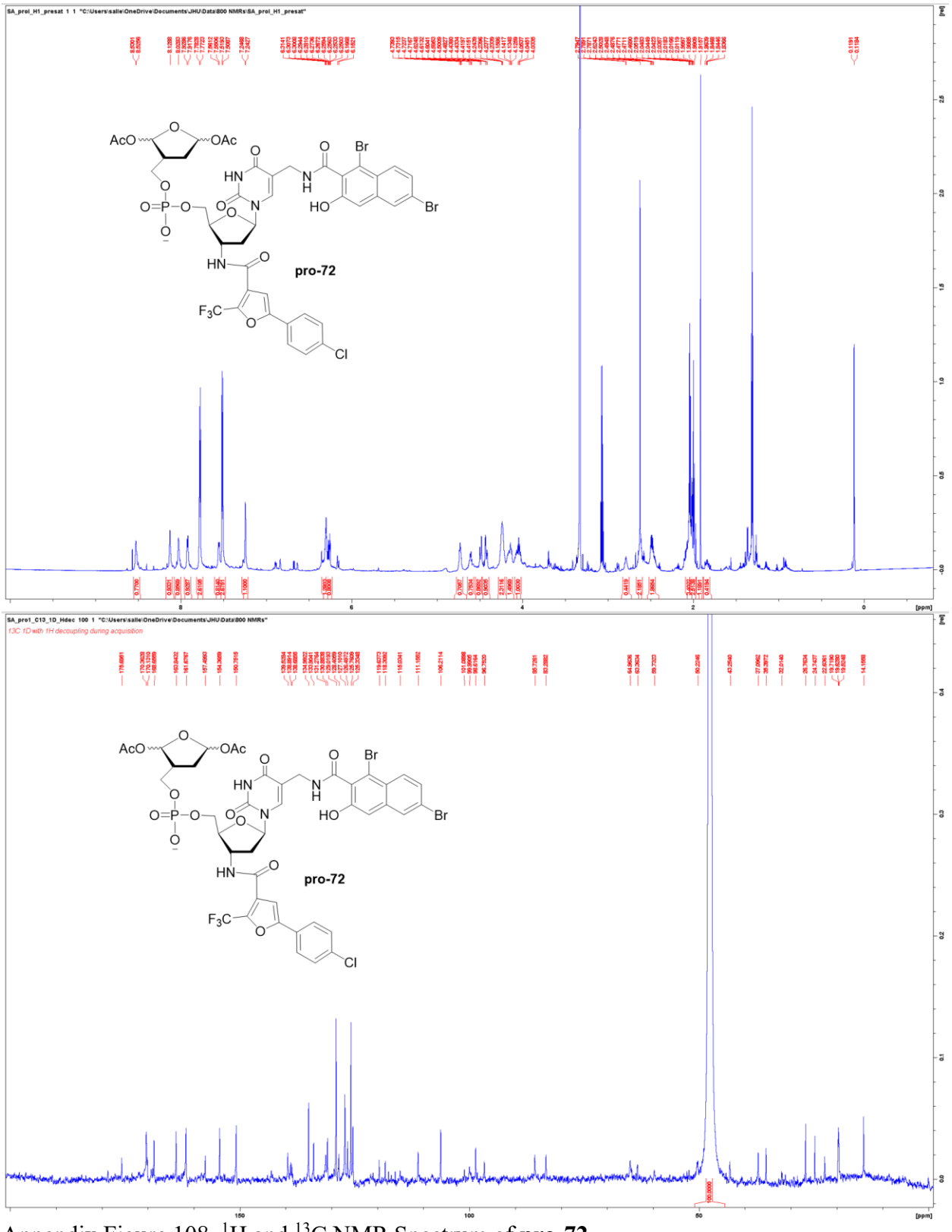
Appendix Figure 105. ^{31}P NMR Spectrum of **74**.



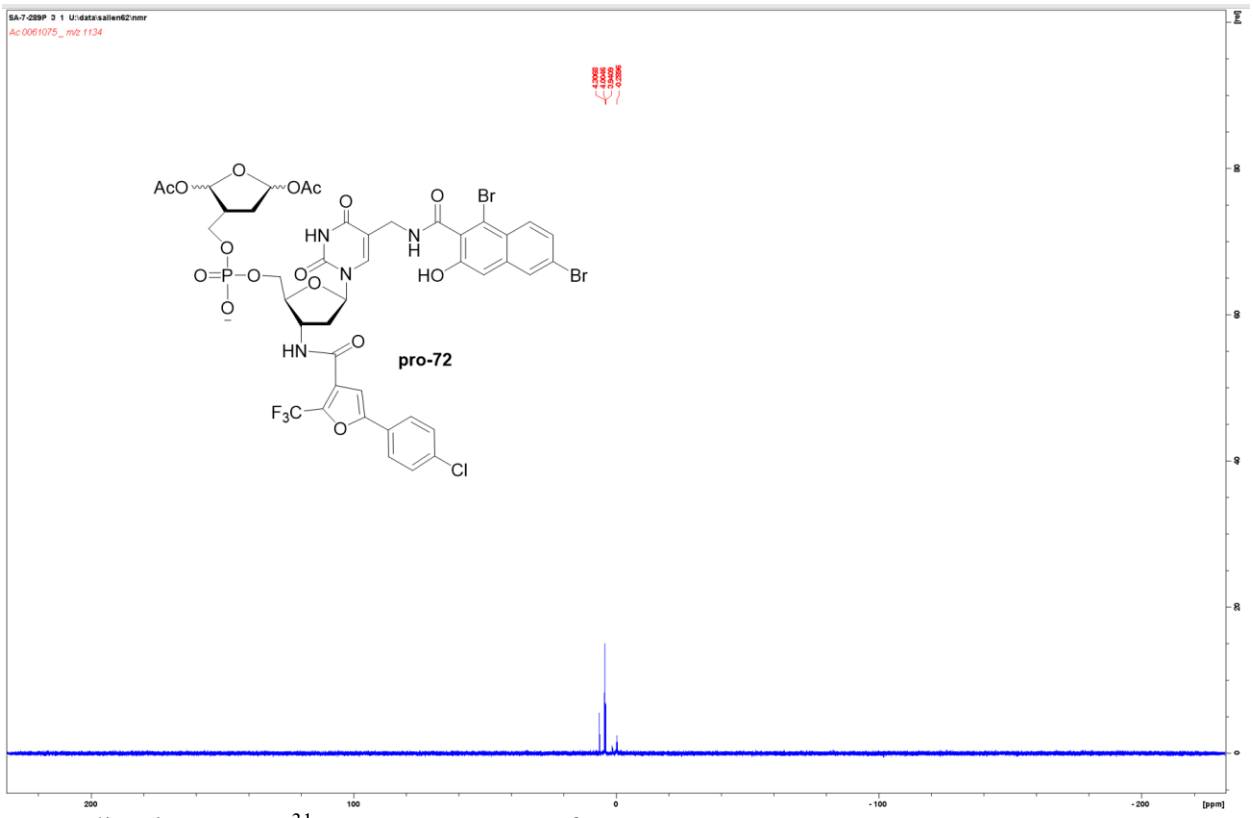
Appendix Figure 106. ^{1}H NMR Spectrum of Precursor to **pro-72**.



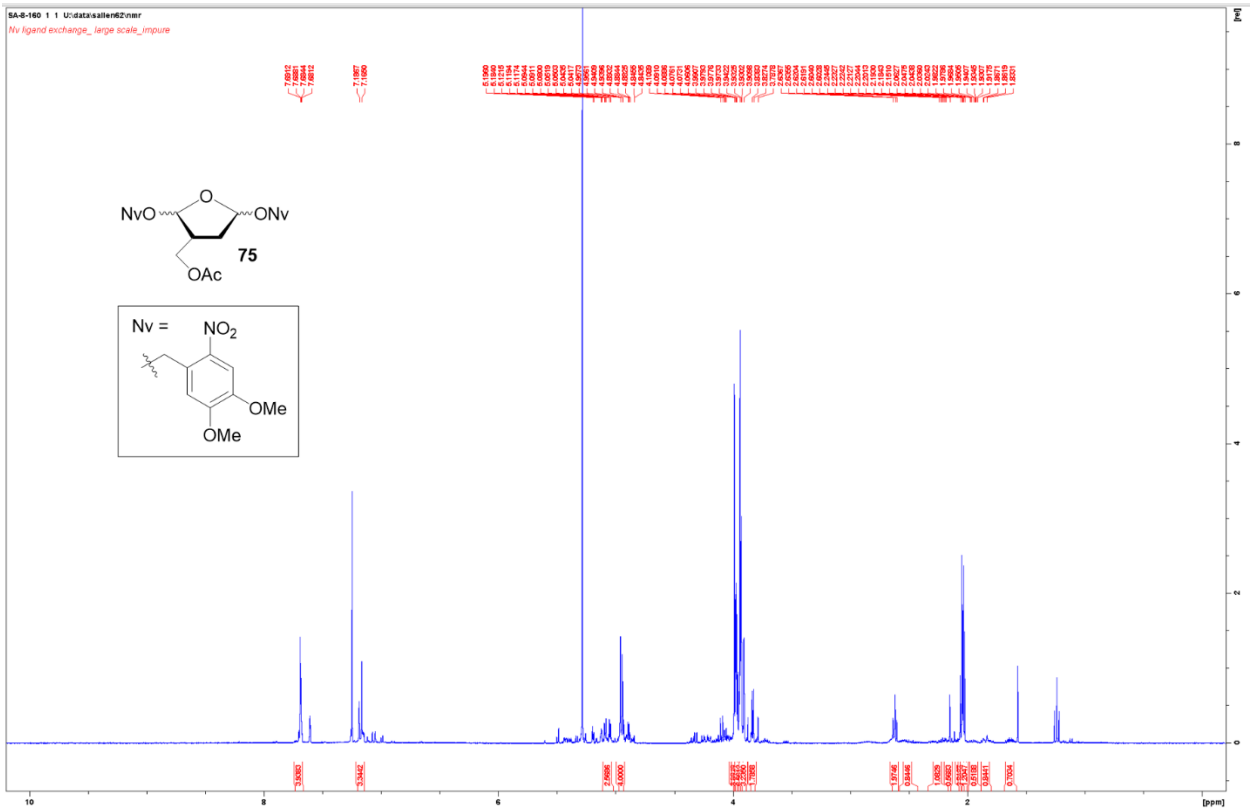
Appendix Figure 107. ^{13}C and ^{31}P NMR Spectrum of Precursor to **pro-72**.



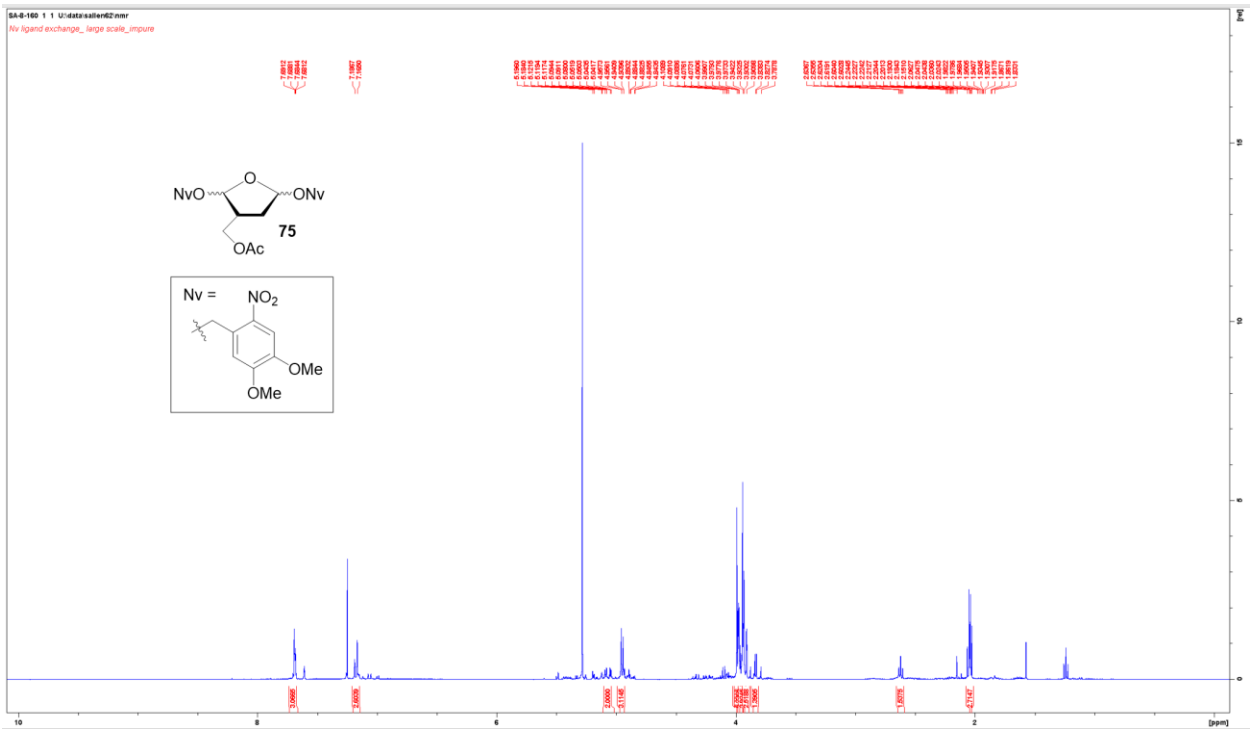
Appendix Figure 108. ^{1}H and ^{13}C NMR Spectrum of **pro-72**.



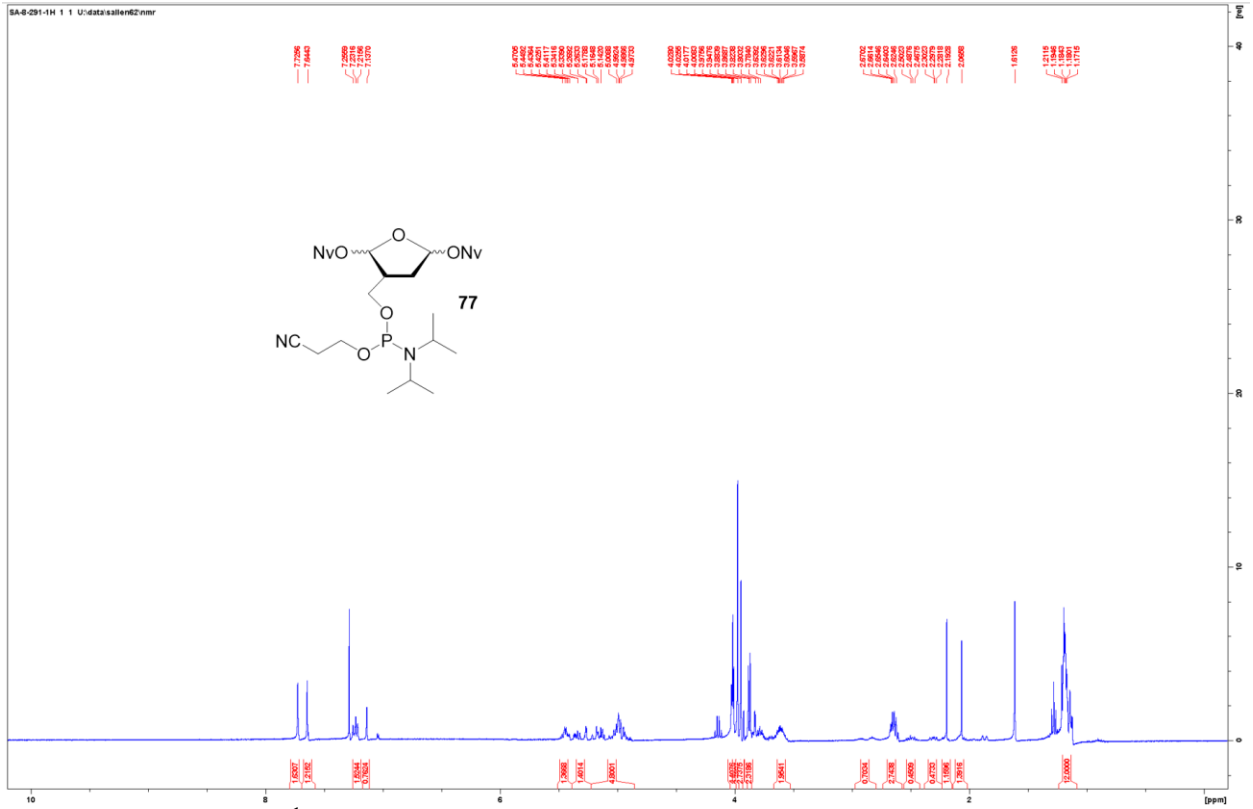
Appendix Figure 109. ^{31}P NMR Spectrum of **pro-72**.



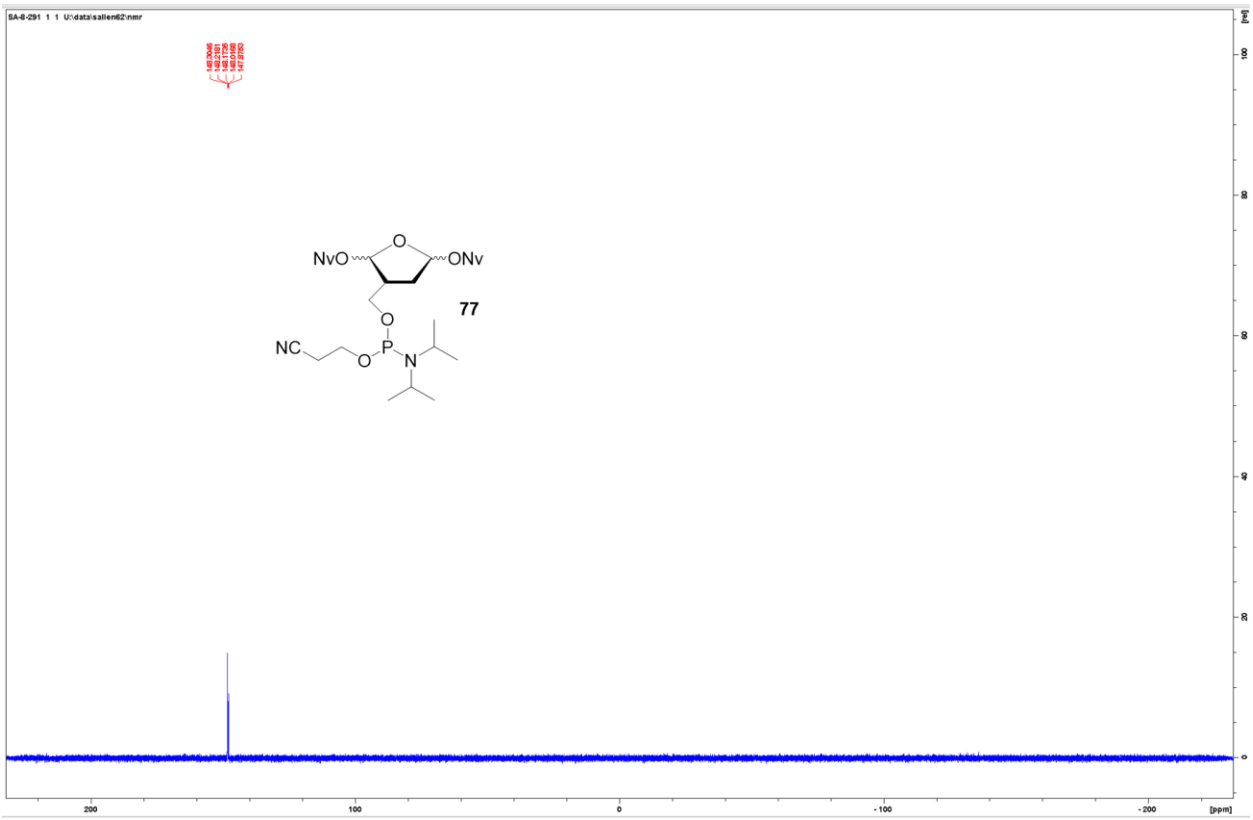
Appendix Figure 110. ^{1}H NMR Spectrum of **75**.



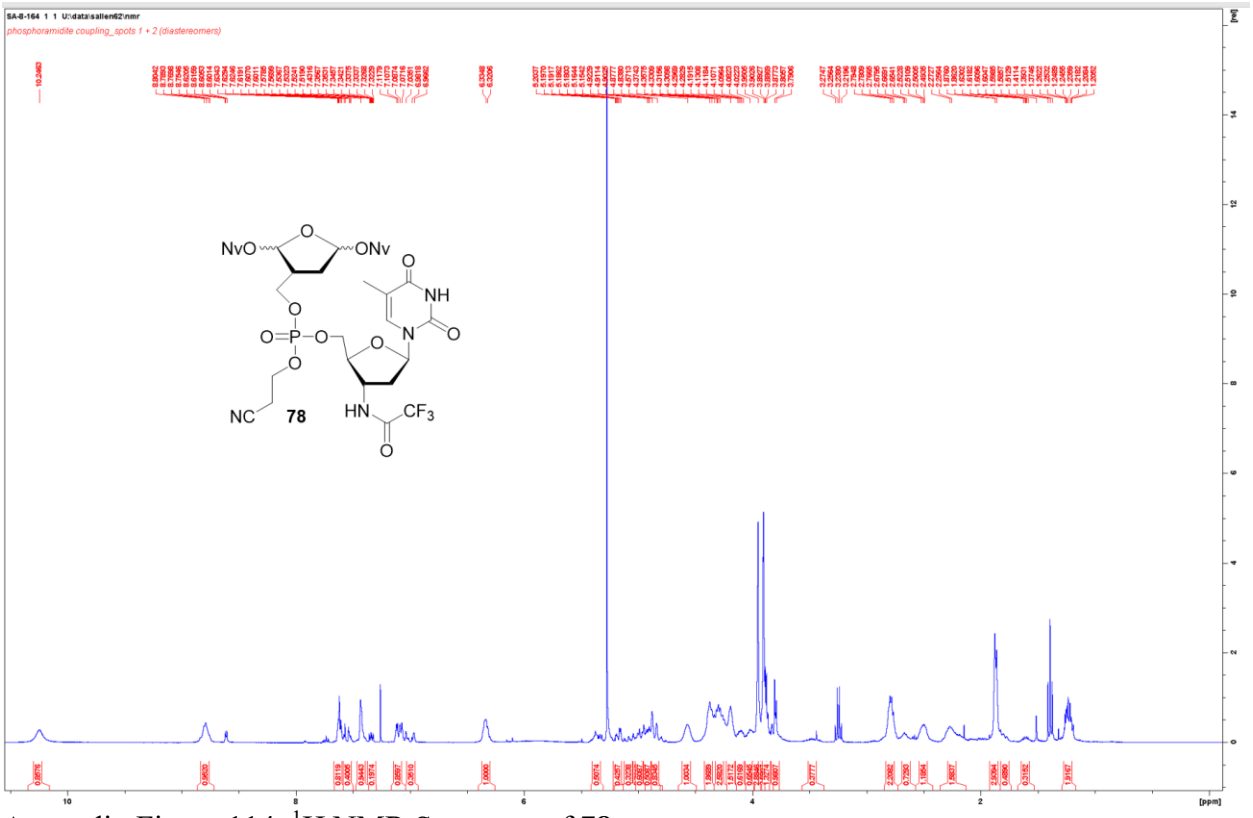
Appendix Figure 111. ^{1}H NMR Spectrum of **75**.



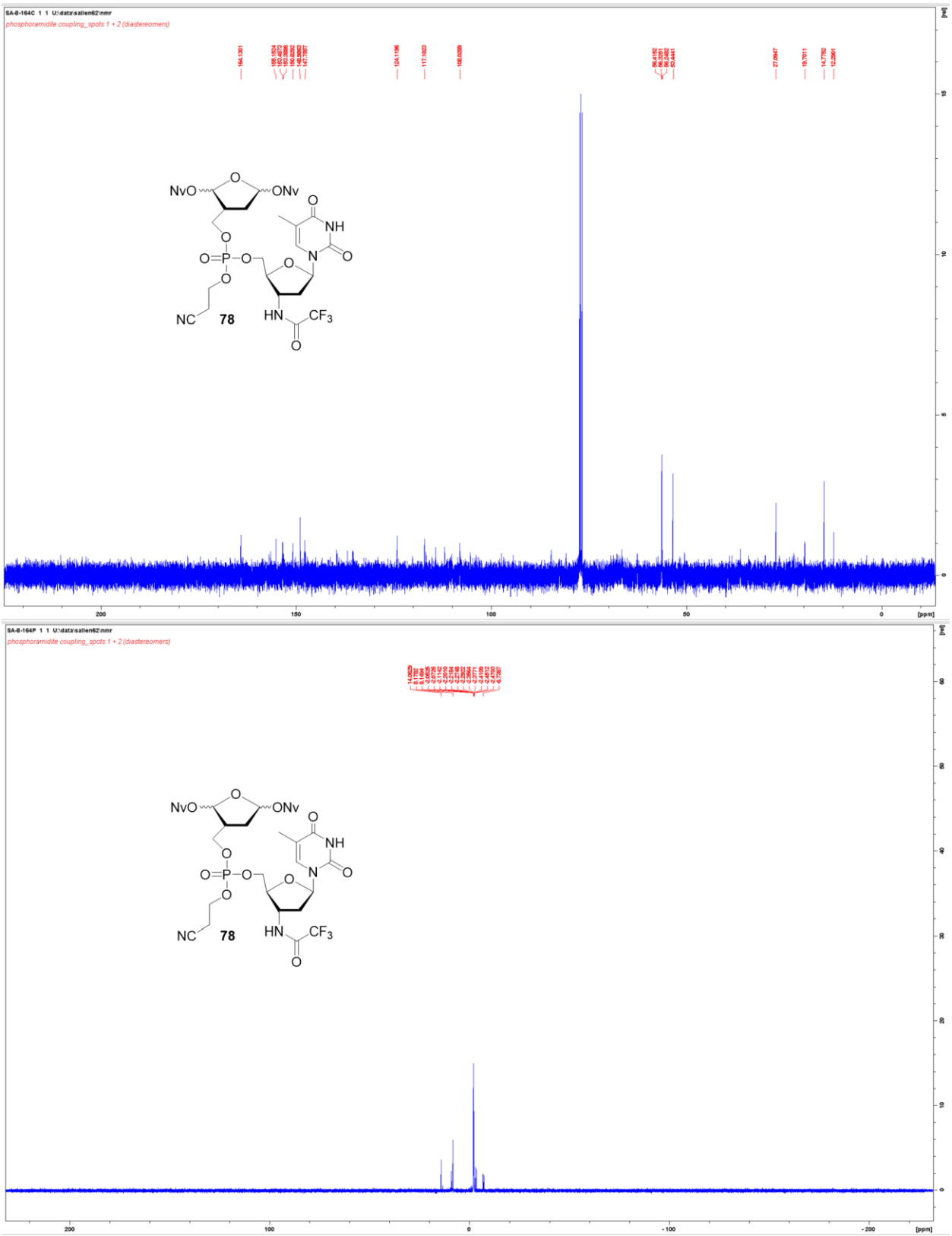
Appendix Figure 112. ^{1}H NMR Spectrum of **77**.



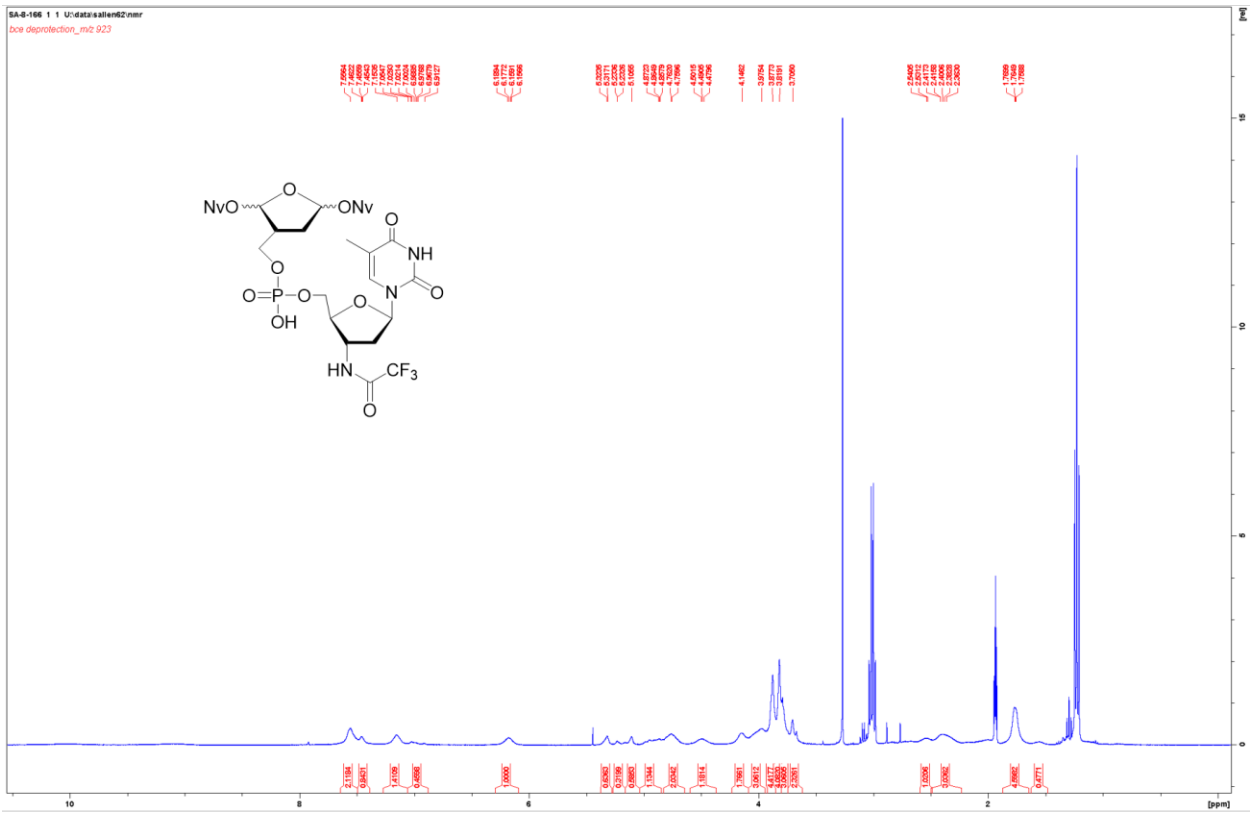
Appendix Figure 113. ^{31}P NMR Spectrum of **77**.



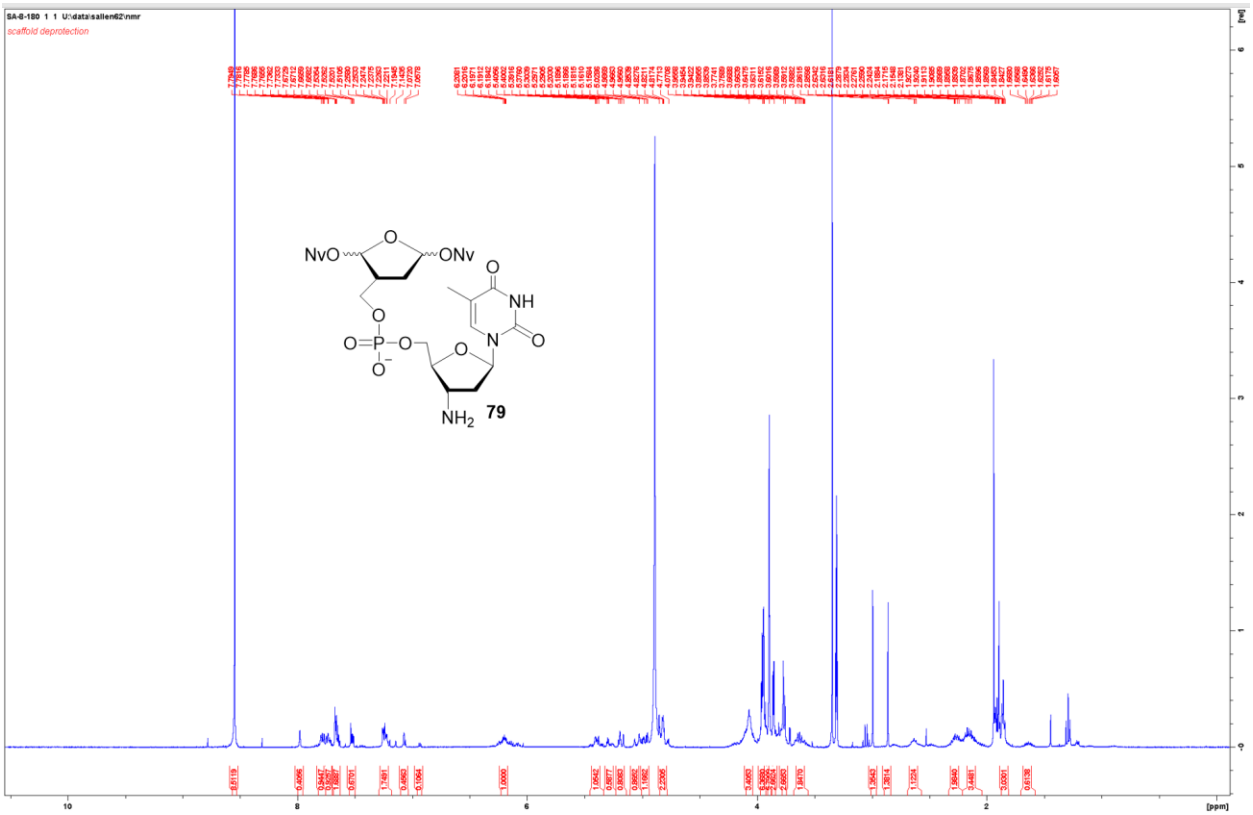
Appendix Figure 114. ^{1}H NMR Spectrum of **78**.



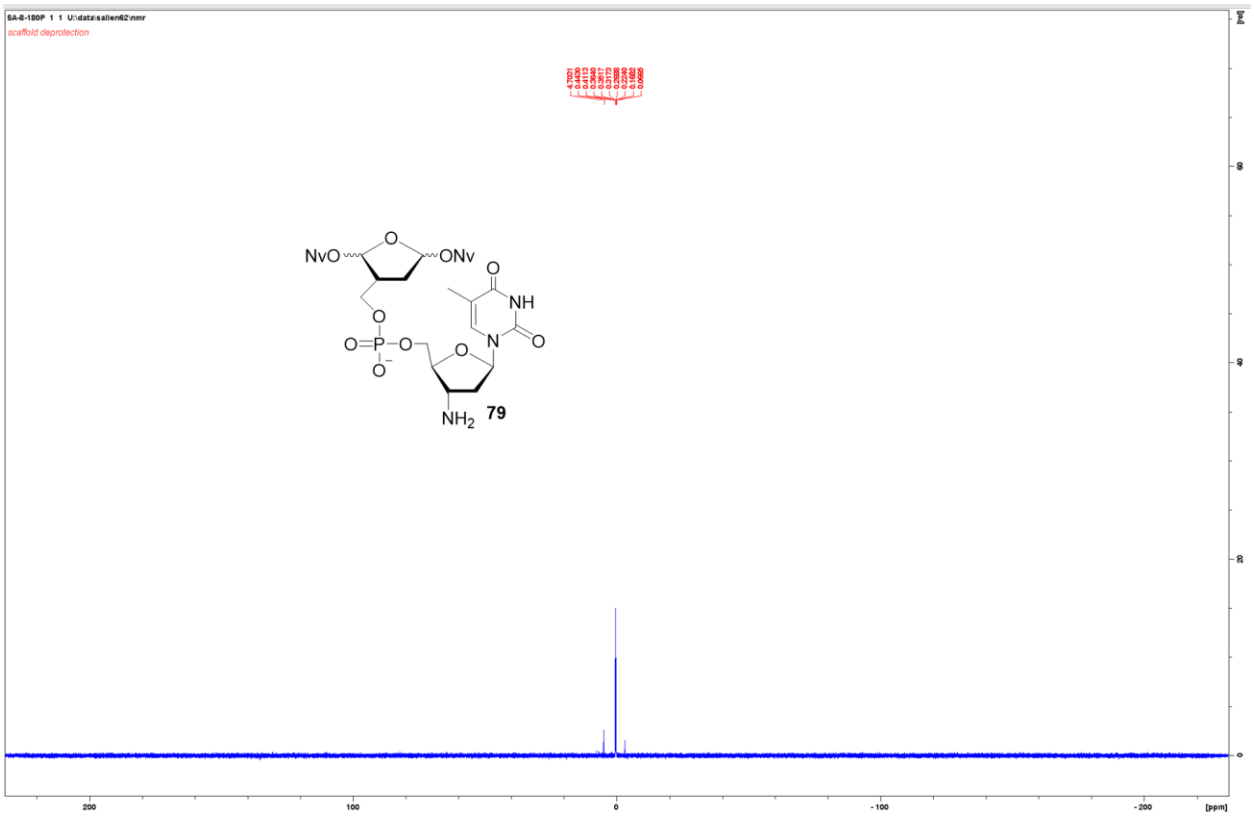
Appendix Figure 115. ^{13}C and ^{31}P NMR Spectrum of **78**.



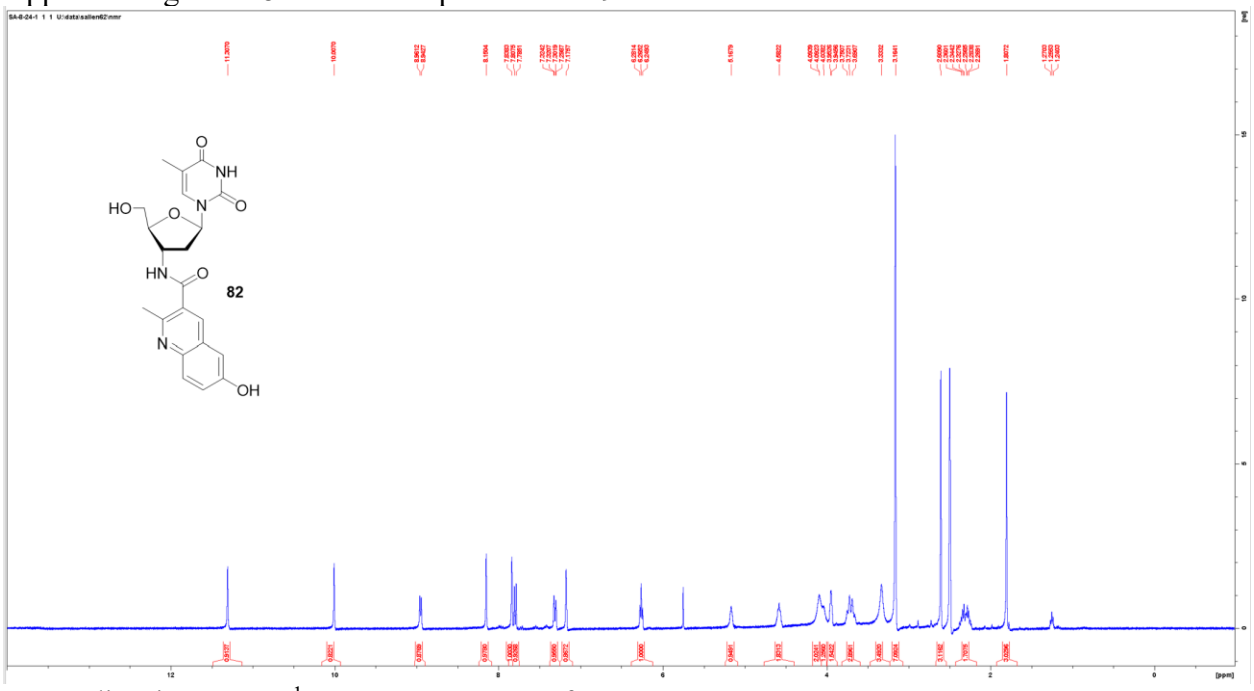
Appendix Figure 116. ^{1}H NMR Spectrum of Precursor to **79**.



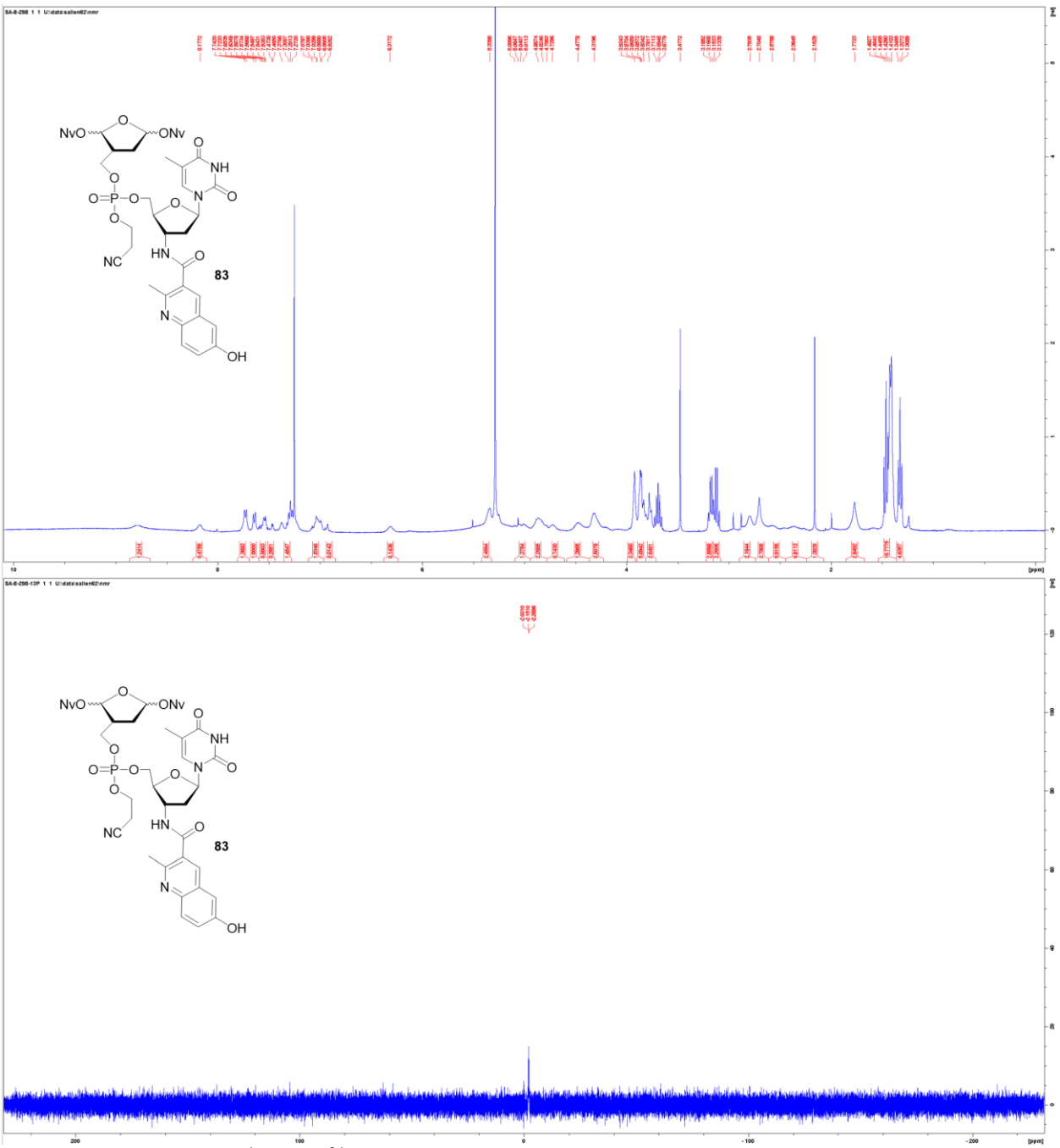
Appendix Figure 117. ^{1}H NMR Spectrum of **79**.



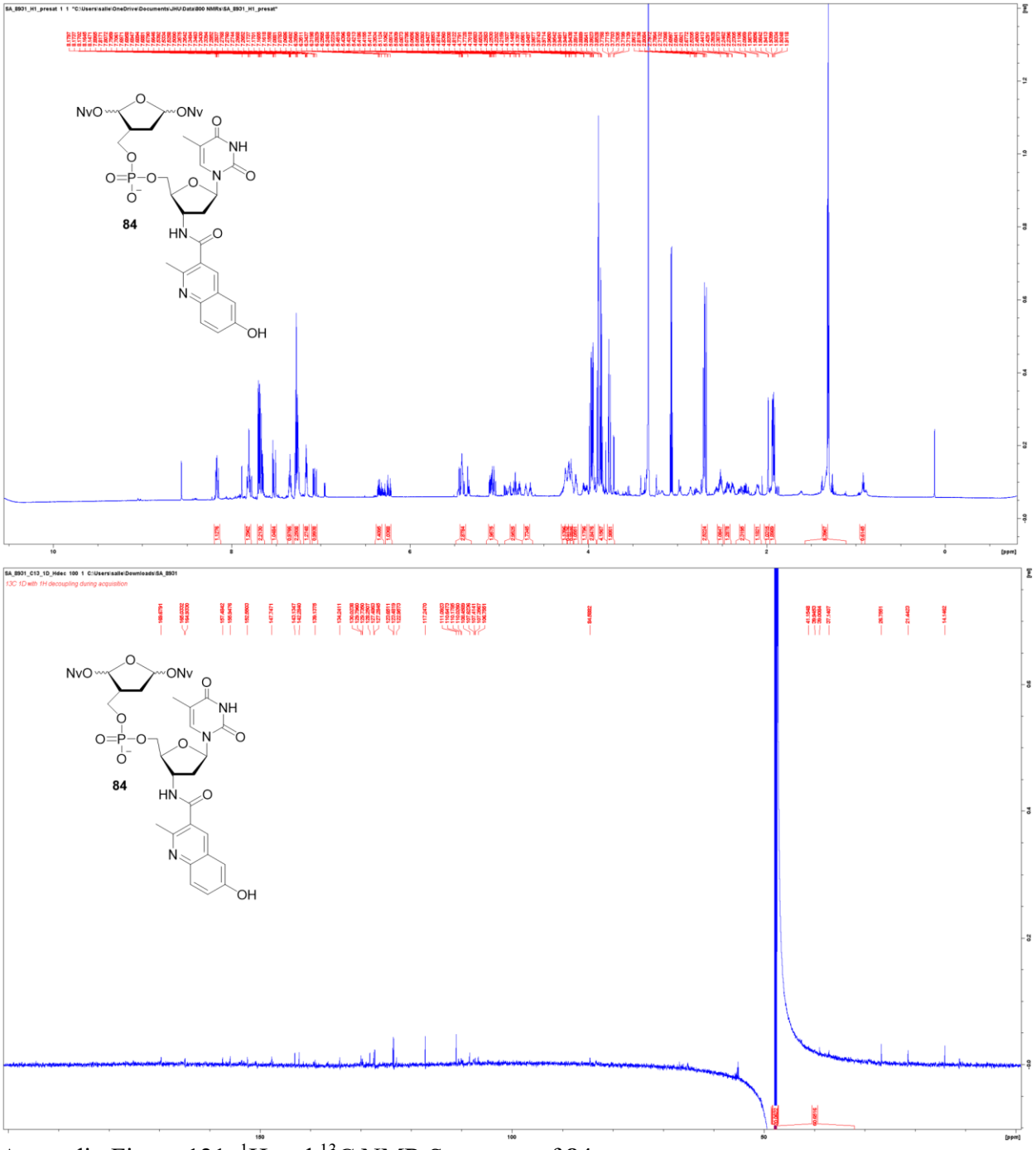
Appendix Figure 118. ^{31}P NMR Spectrum of **79**.



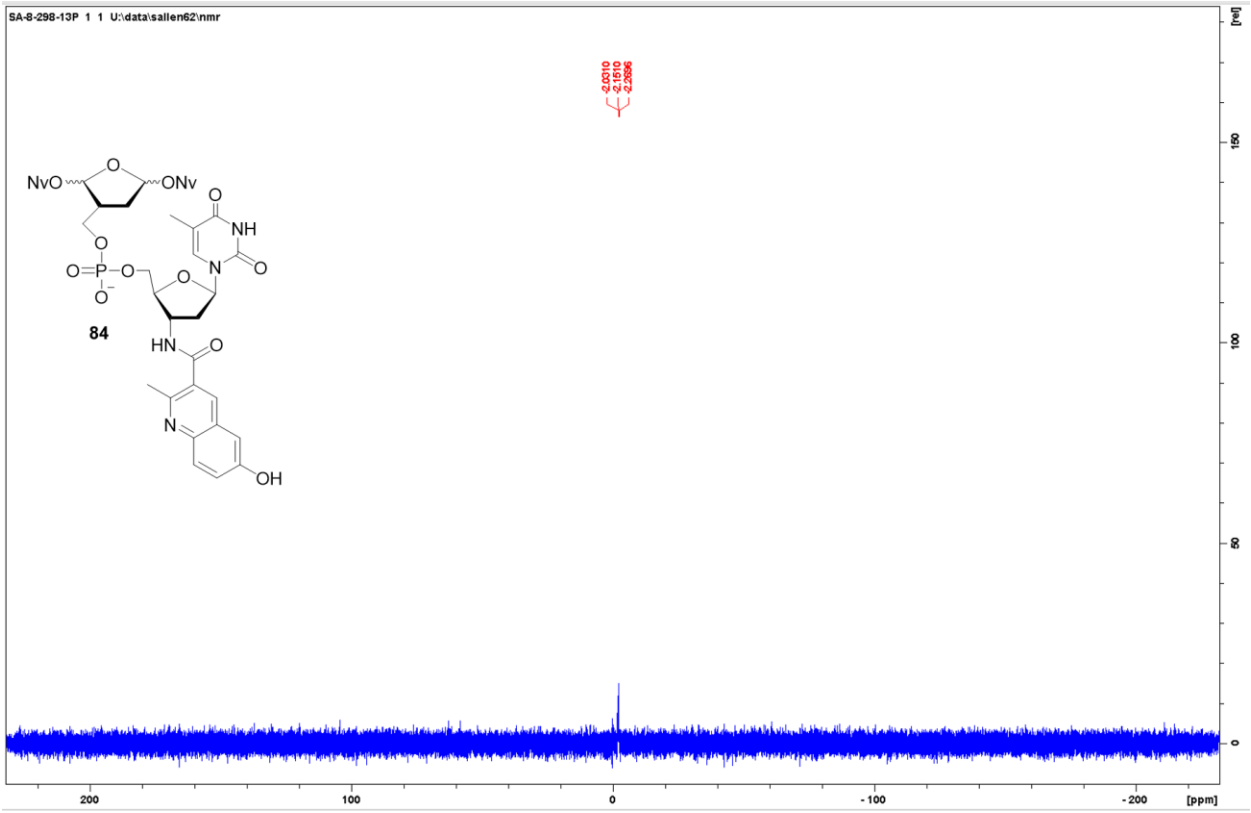
Appendix Figure 119. ^{1}H NMR Spectrum of **82.**



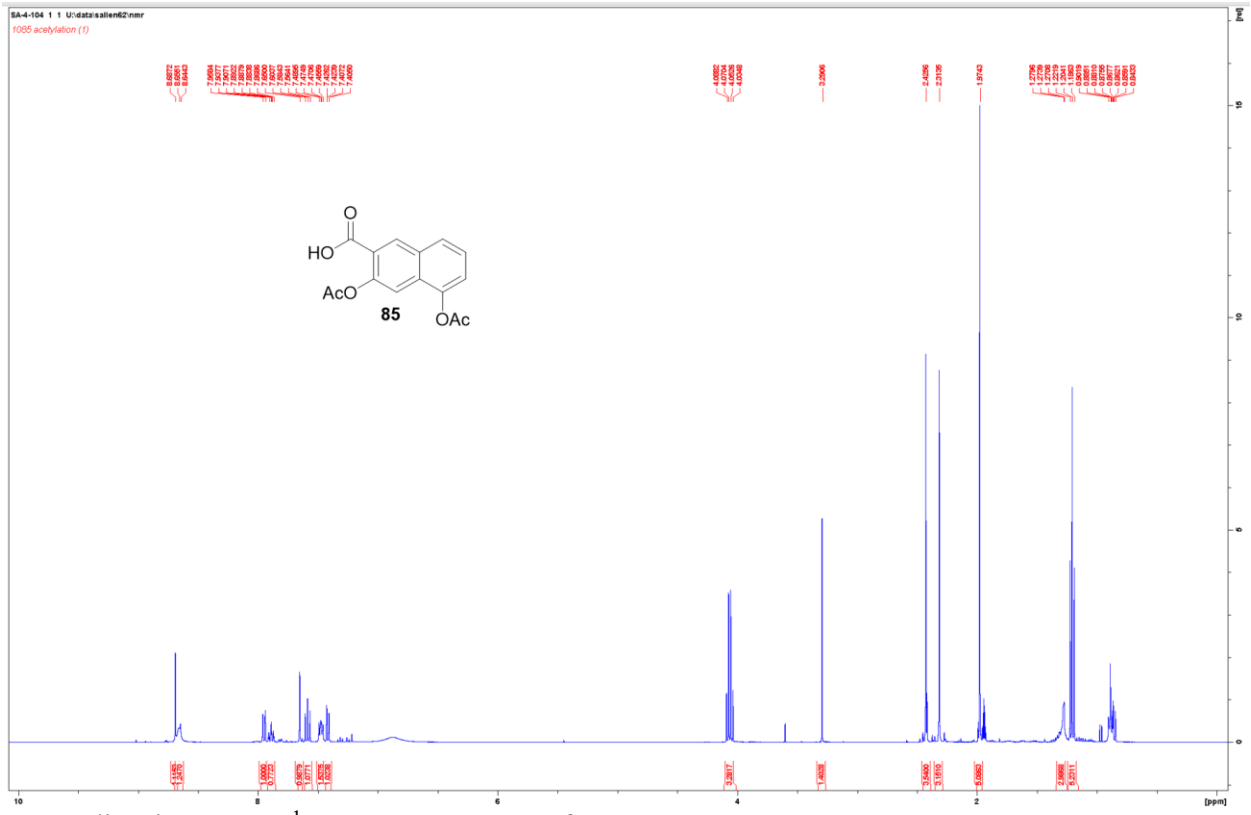
Appendix Figure 120. ^{1}H and ^{31}P NMR Spectrum of **83.**



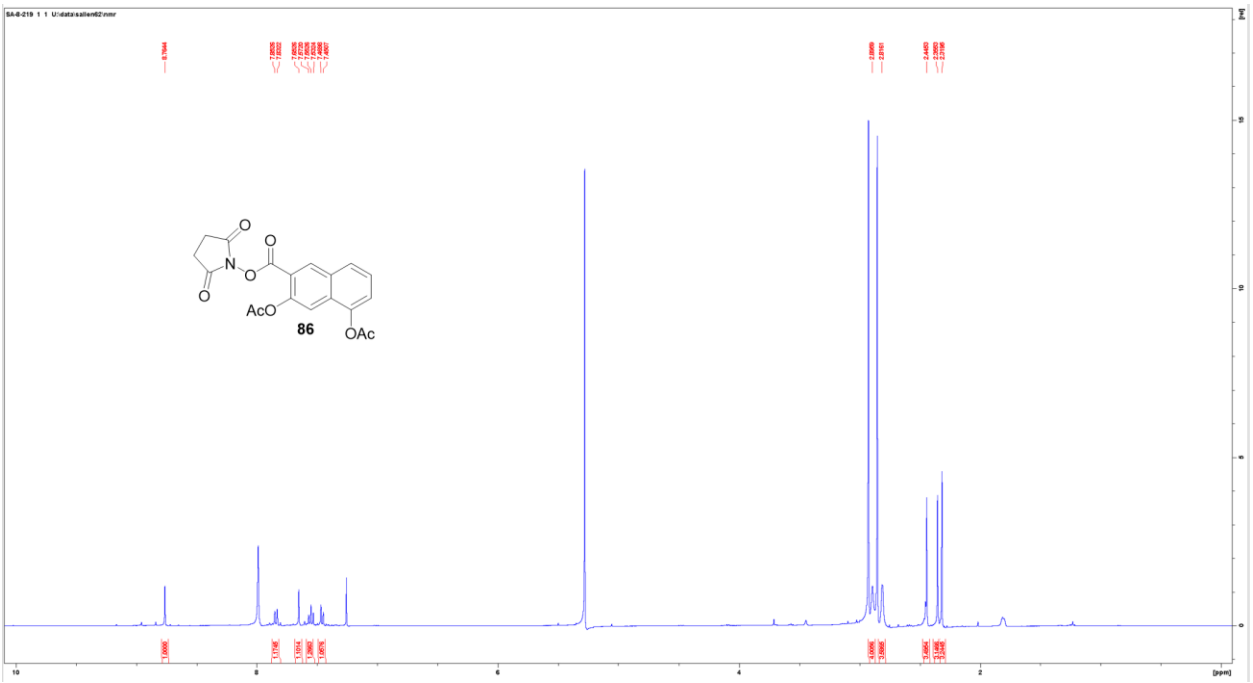
Appendix Figure 121. ^{1}H and ^{13}C NMR Spectrum of **84**.



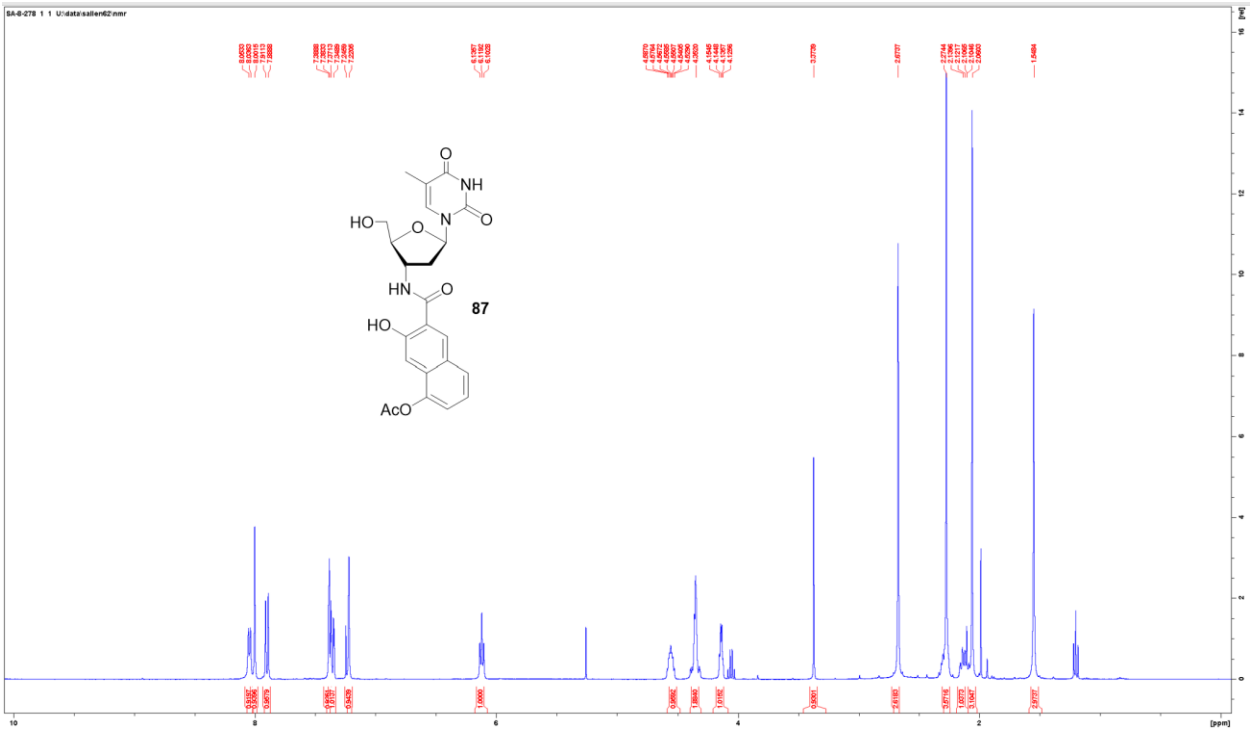
Appendix Figure 122. ^{31}P NMR Spectrum of **84.**



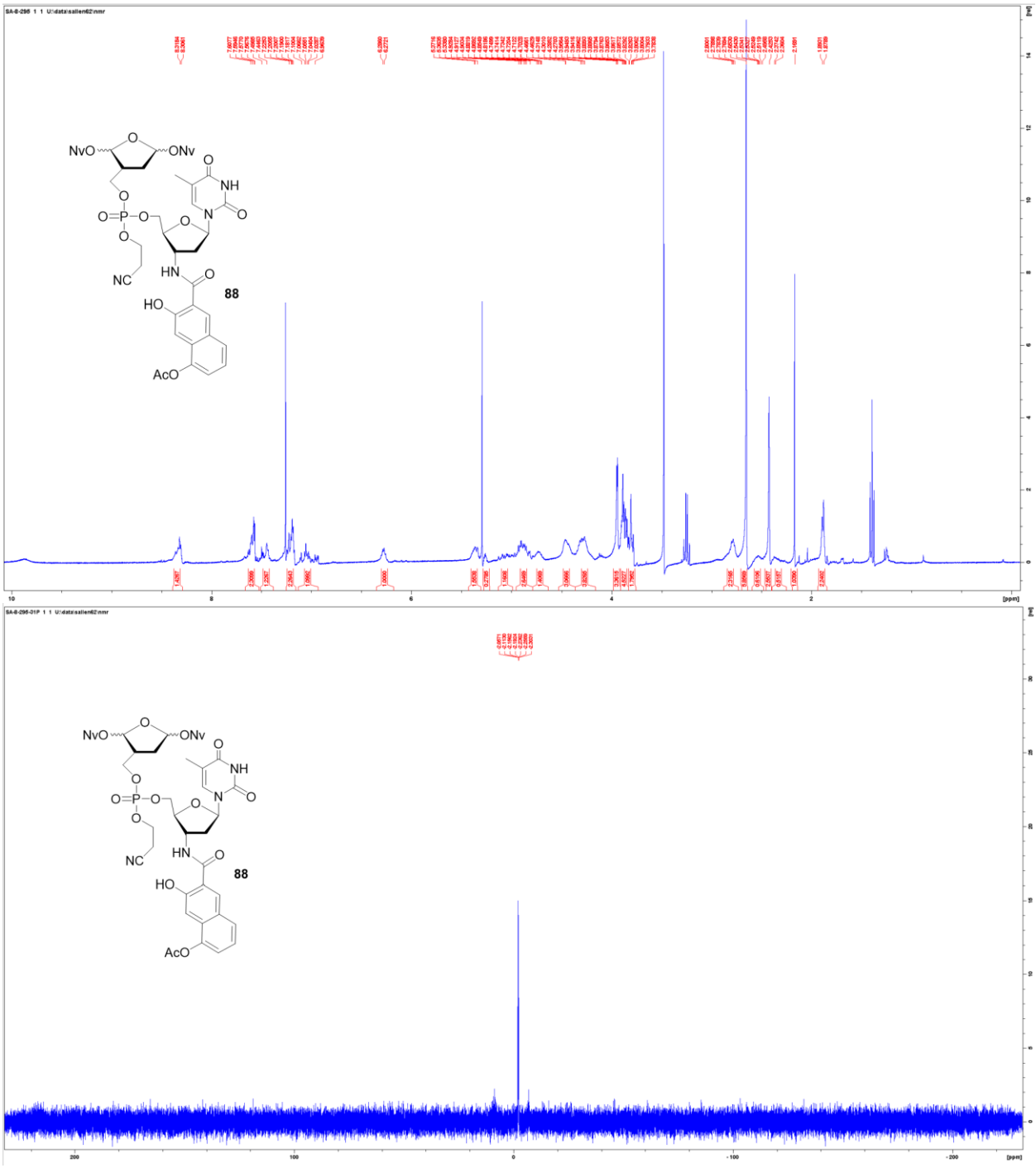
Appendix Figure 123. ^{1}H NMR Spectrum of **85.**



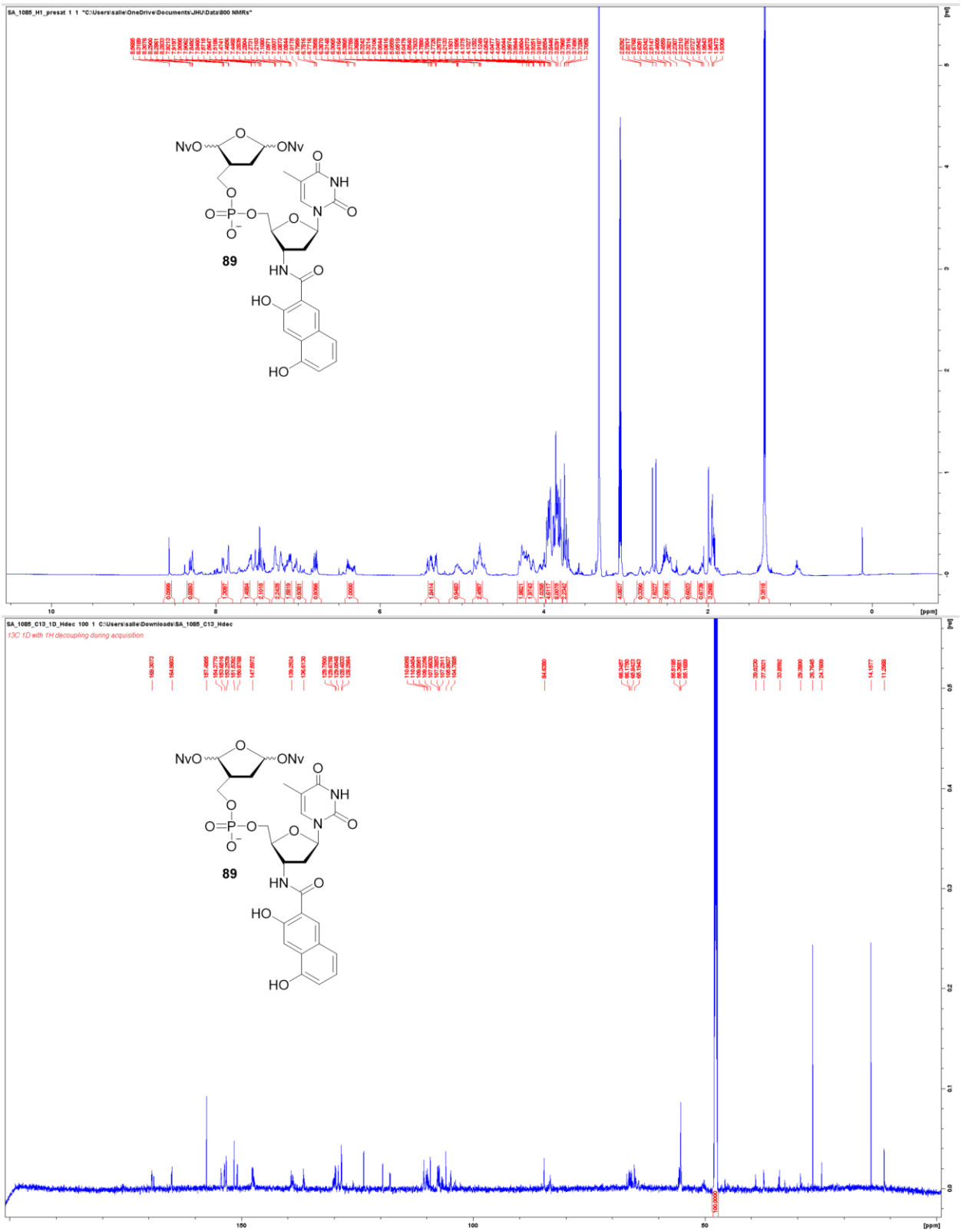
Appendix Figure 124. ^{1}H NMR Spectrum of **86.**



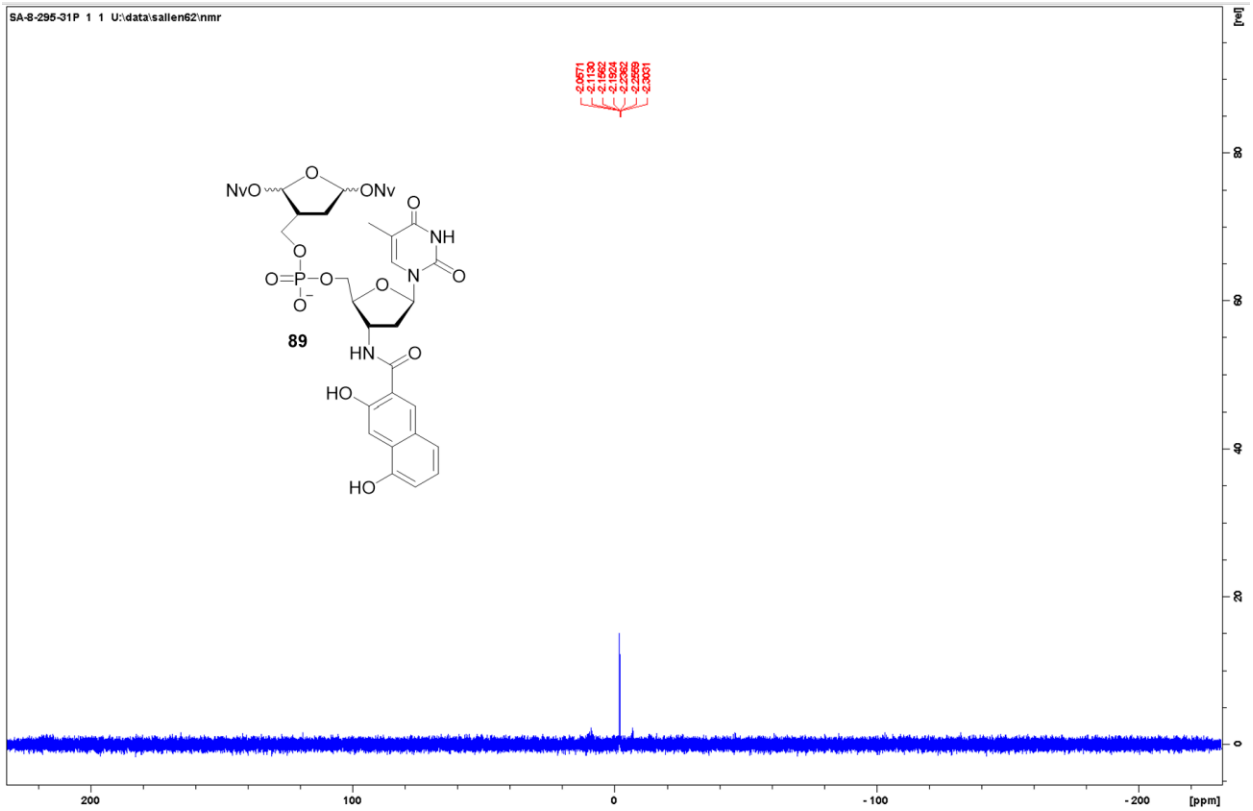
Appendix Figure 125. ^{1}H NMR Spectrum of **87.**



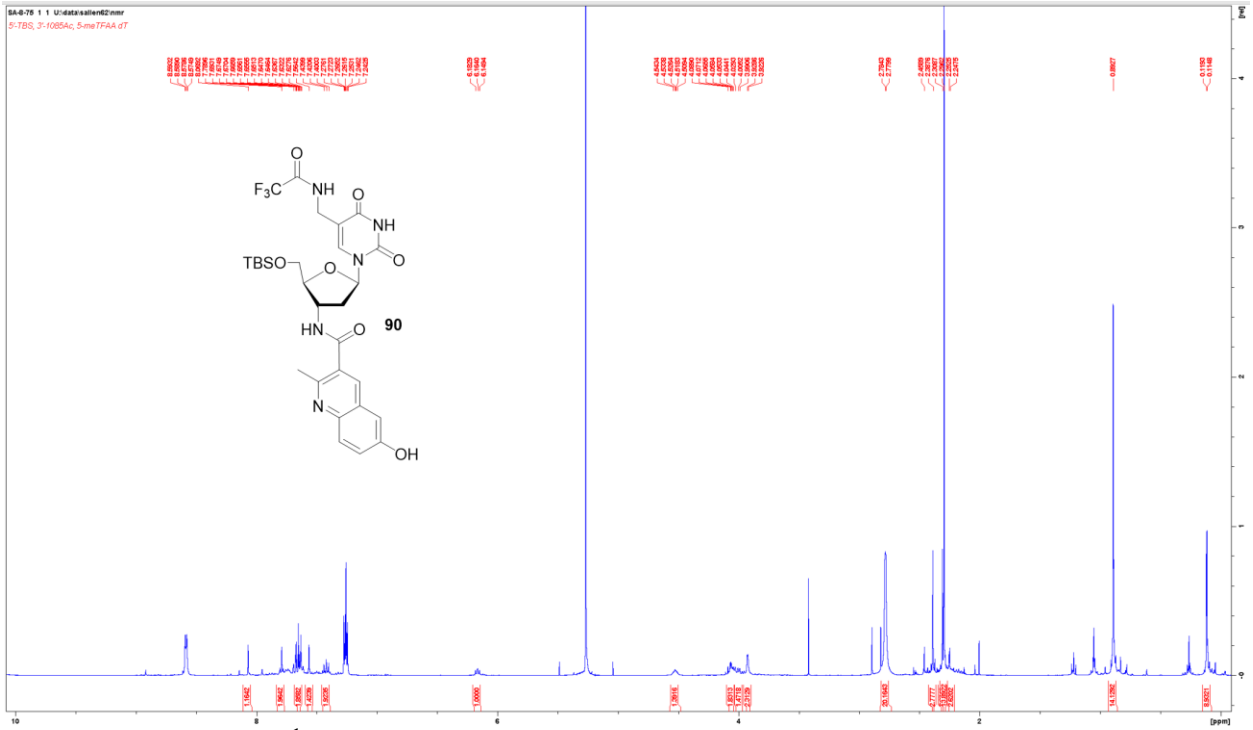
Appendix Figure 126. ^{1}H and ^{31}P NMR Spectrum of **88.**



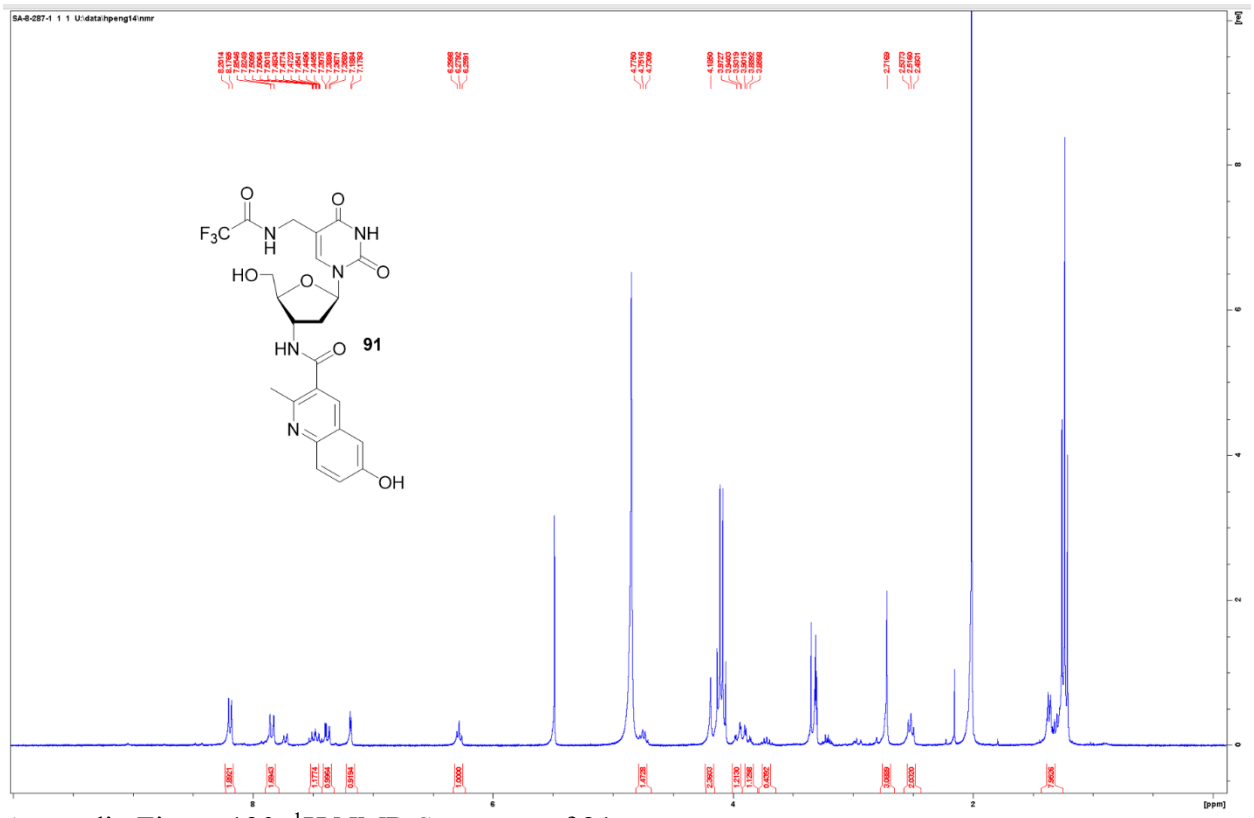
Appendix Figure 127. ^{1}H and ^{13}C NMR Spectrum of **89.**



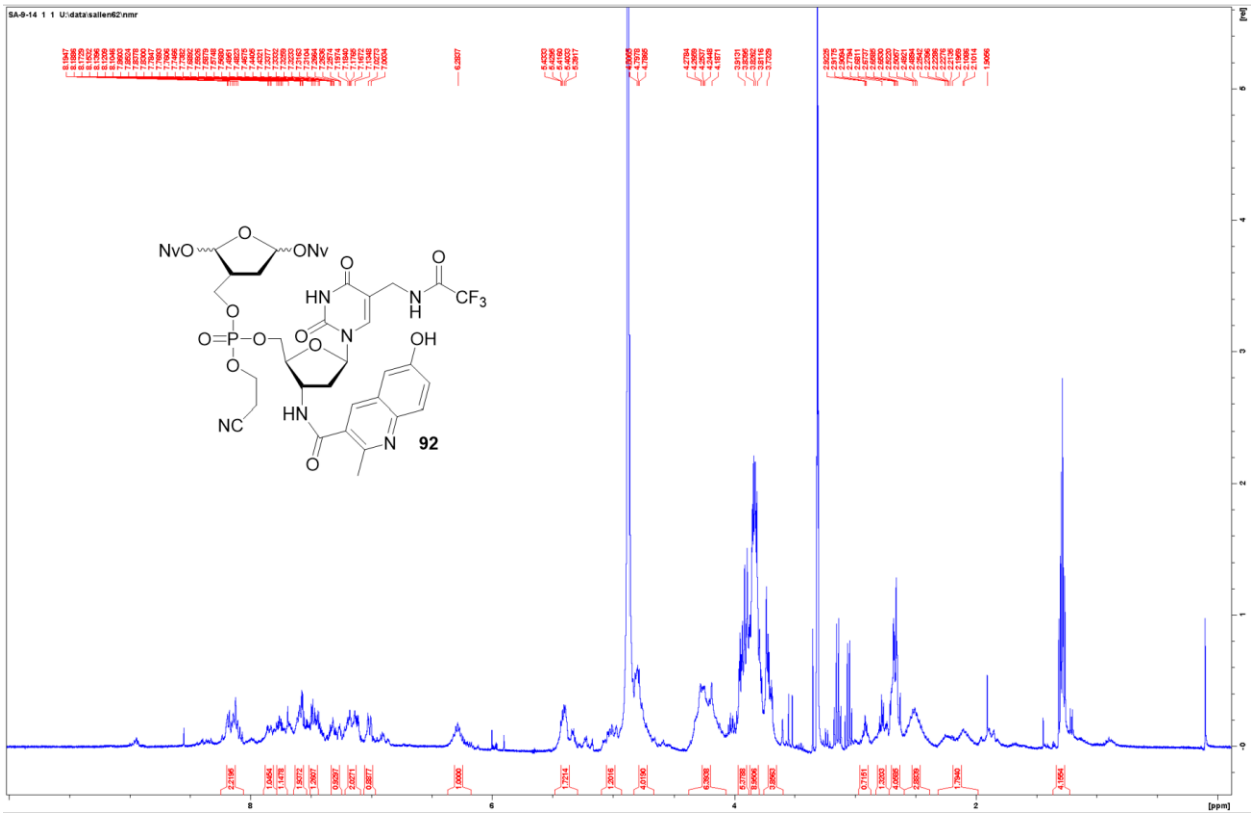
Appendix Figure 128. ^{31}P NMR Spectrum of **89.**



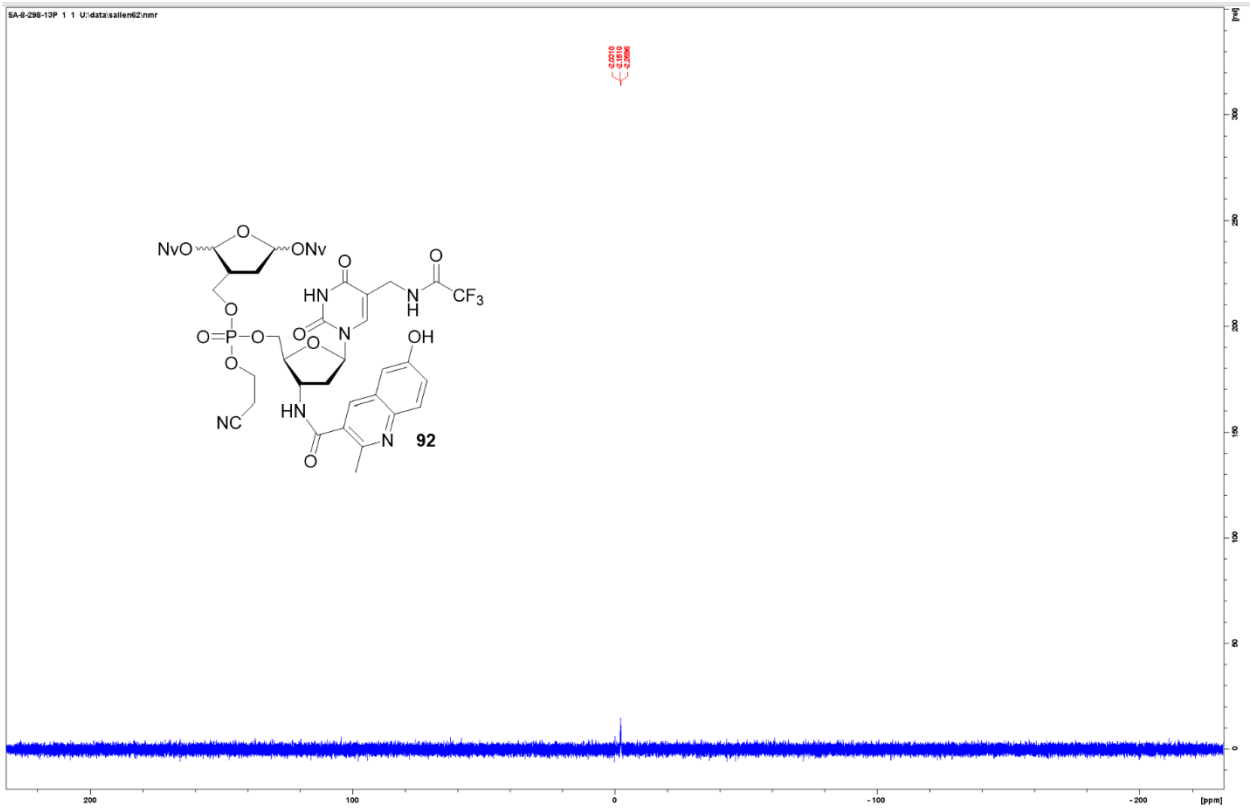
Appendix Figure 129. ^{1}H NMR Spectrum of **90.**



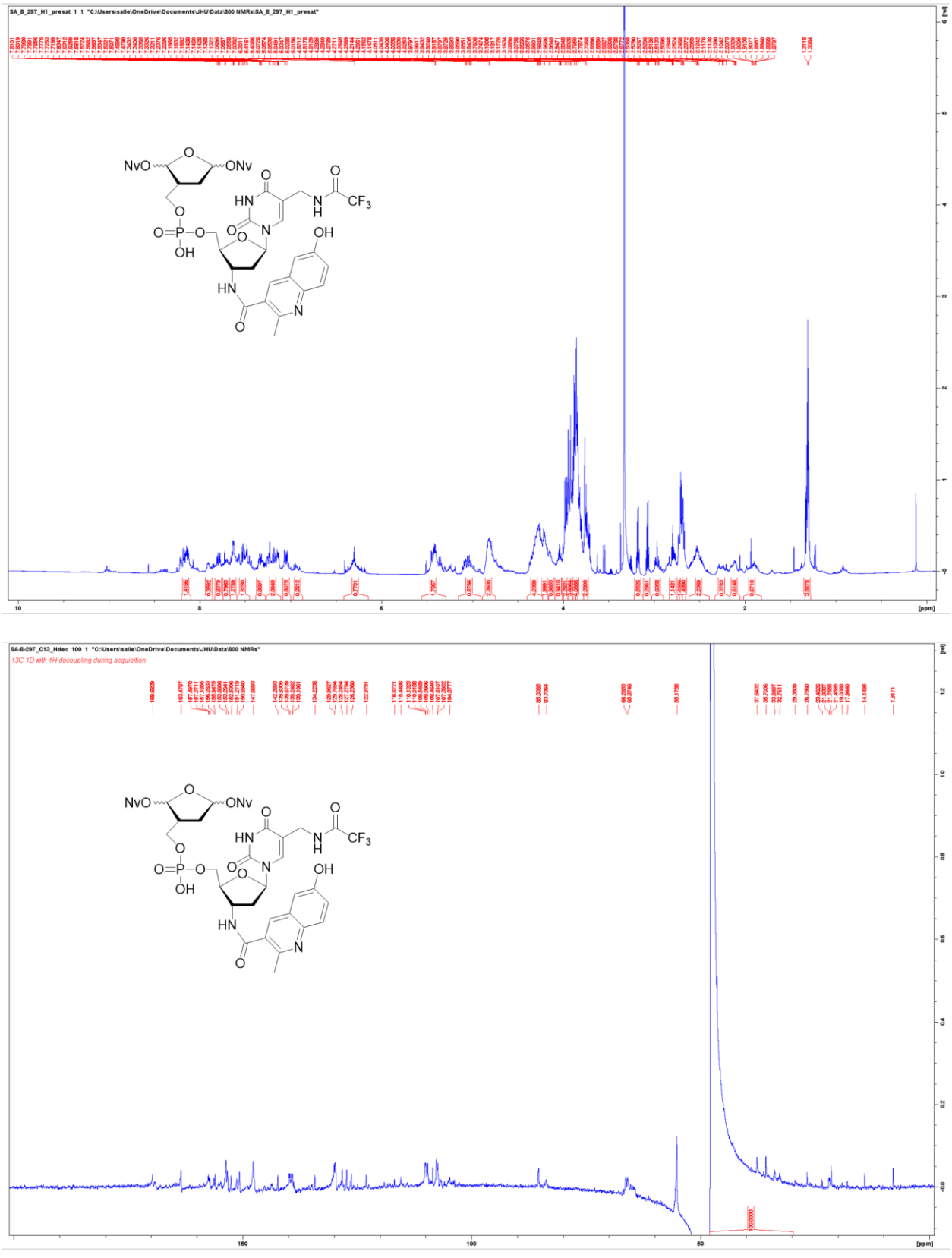
Appendix Figure 130. ^{1}H NMR Spectrum of **91**.



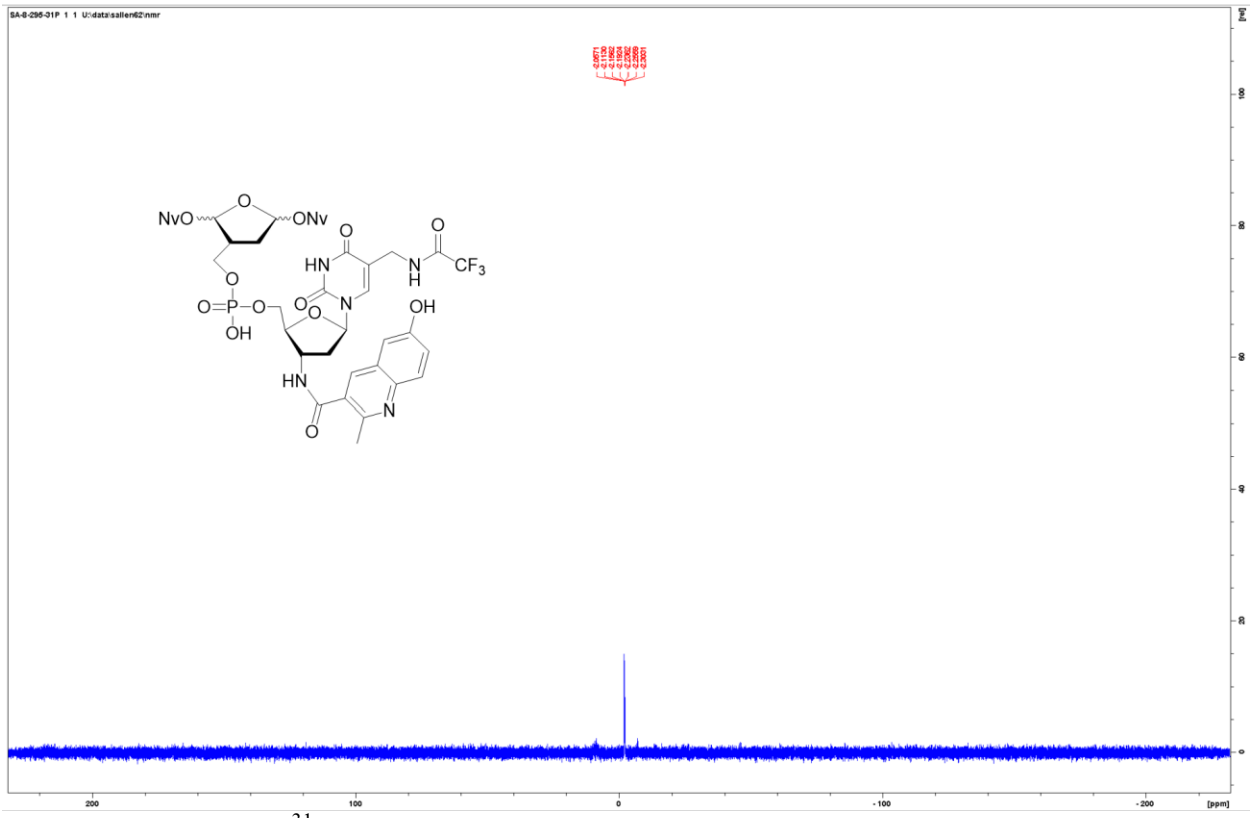
Appendix Figure 131. ^{1}H NMR Spectrum of **92.**



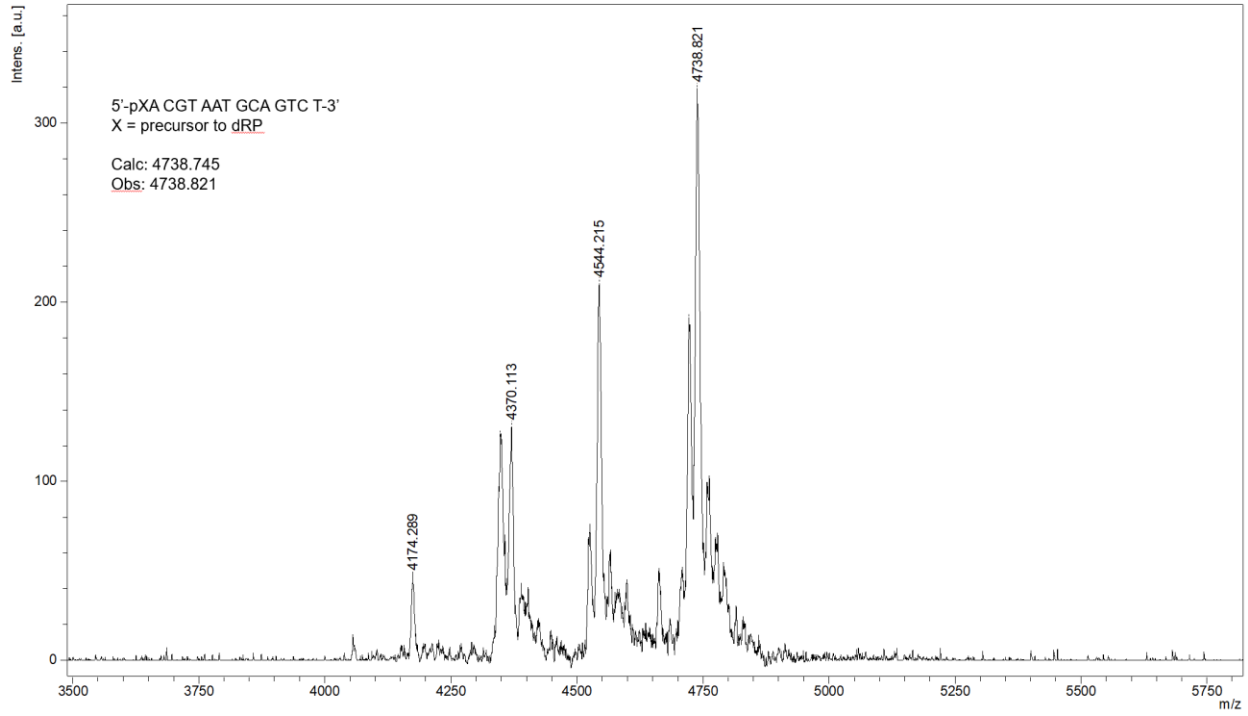
Appendix Figure 132. ^{31}P NMR Spectrum of **92.**



Appendix Figure 133. ^{1}H and ^{13}C NMR Spectrum of Precursor to **93.**



Appendix Figure 134. ^{31}P NMR Spectrum of Precursor to **93.**



Appendix Figure 135. MALDI-TOF of 5'-dRP precursor oligonucleotide; m/z calculated for (M + H)− 4738.745, 4738.821 observed.

Appendix 3. 1D and 2D NMR Analysis of final inhibitors.

Analysis of **pro-72:**

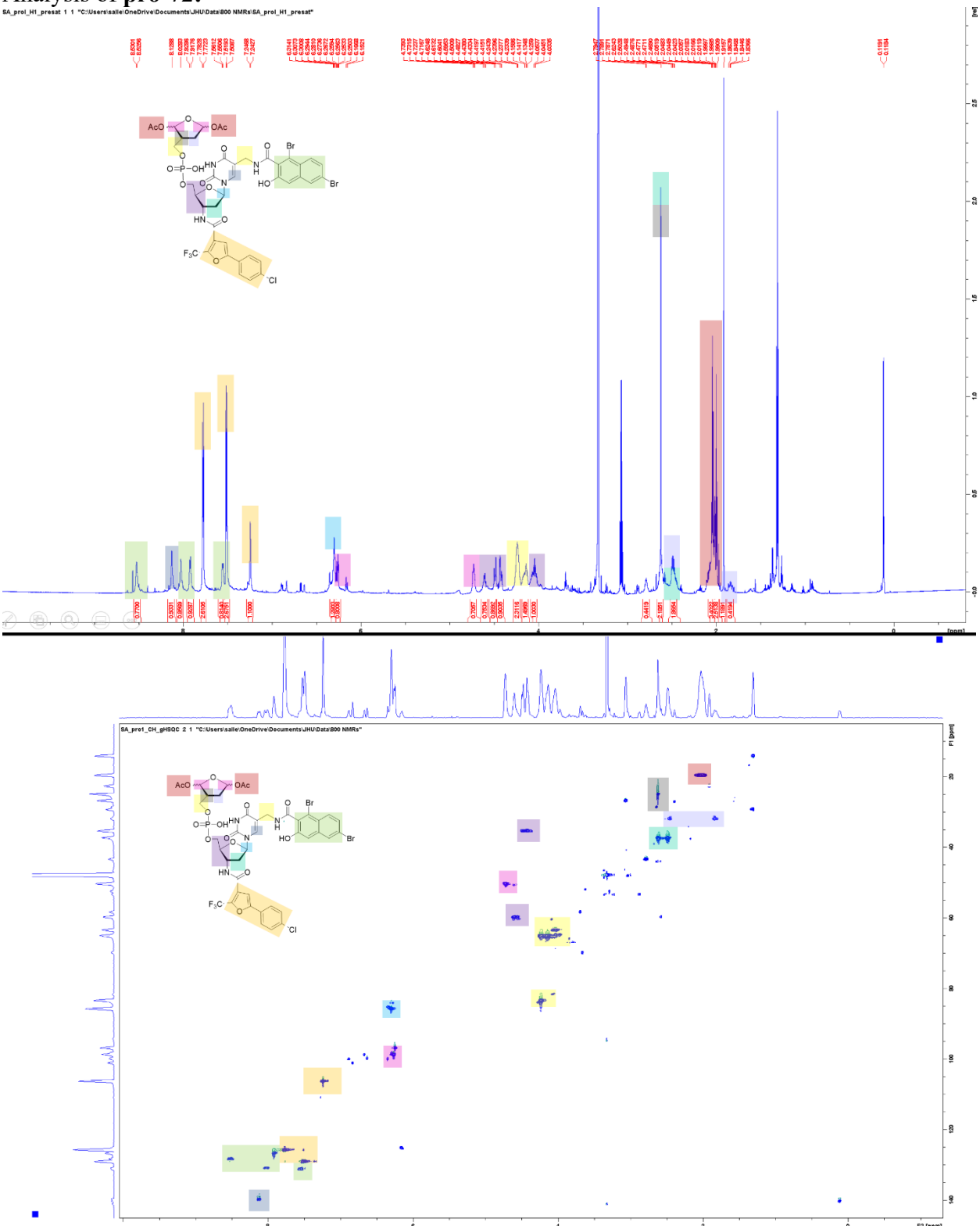


Table 24. Chemical Shift Analysis of **pro-72** NMR spectra.

Proton			Carbon	
Chemical Shift	**Splitting (Coupling Constant)**	**Integration**	**Chemical Shift**	**Description**
8.53	s	1	128.4	1075 (1/4)
8.12	S	1	139.5	dT aromatic (1/1)
8.02	S	1	130.8	1075 (2/4)
7.92	D (J = 8.8 Hz)	1	126.5	1075 (3/4
7.78	D (J = 8.5 Hz)	2	125.7	006 aromatic (2/2)
7.56	D (J = 8.8 Hz)	1	131.3	1075 (4/4)
7.53	D (J = 8.5 Hz)	2	129.0	006 aromatic (2/2)
7.24	S	1	106.2	006 furan (1/1)
6.30	Quint (J = 5.4 Hz)	1	85.7	C1' (1/1) (~85 ppm)
6.25	M	1	96.7	Acetal (1/2) (~95-110 ppm)
4.73	T (J = 6 Hz)	1	50.2	Acetal (2/2)
4.61	Q (J = 7 Hz)	1	59.7	C3' (1/1)
4.49	D (J = 14 Hz)	1	35.4	C5' (1/2)
4.42	Dd (J = 14 Hz, 3.6 Hz)	1	35.4	C5' (2/2)
4.24	M	3	65.0	Amide methylene (1/2) (CNR ~45-65 ppm)
			83.3	DOB linker (2/2) (COH ~50-90 ppm)
4.14	M	1	65.0	Amide methylene (2/2)
4.05	T (J = 7 Hz)	1	63.4	C4' (1/1)
2.62	M	2	24.7	DOB methine
			37.1	C2' (1/2)
2.48	Q (J = 6 Hz)	2	37.1	C2' (2/2)
			32.0	DOB C2' (1/2)
2.05	M	3	19.6	Ac (3/3)
1.99	S	3	19.7	Ac (3/3)
1.92	S	1		
1.84	M	1	32.0	DOB C2' (2/2)
			14.2	
			19.5	
			22.6	
			26.7	
			43.2	

	98.6	
	100.0	Acetal CF_3
	101.1	
	111.1	
	115.0	
	118.3	
	119.6	
	125.3	
	127.1	
	133.9	
	135.0	
	138.6	
	150.7	Quaternary C: 006 (4) 1075 (6) Amides (2) dT (1, 1)
	154.3	
	157.5	
	161.7	
	163.8	
	168.6	
	170.1	
	170.4	
	175.69	

Analysis of **84**:

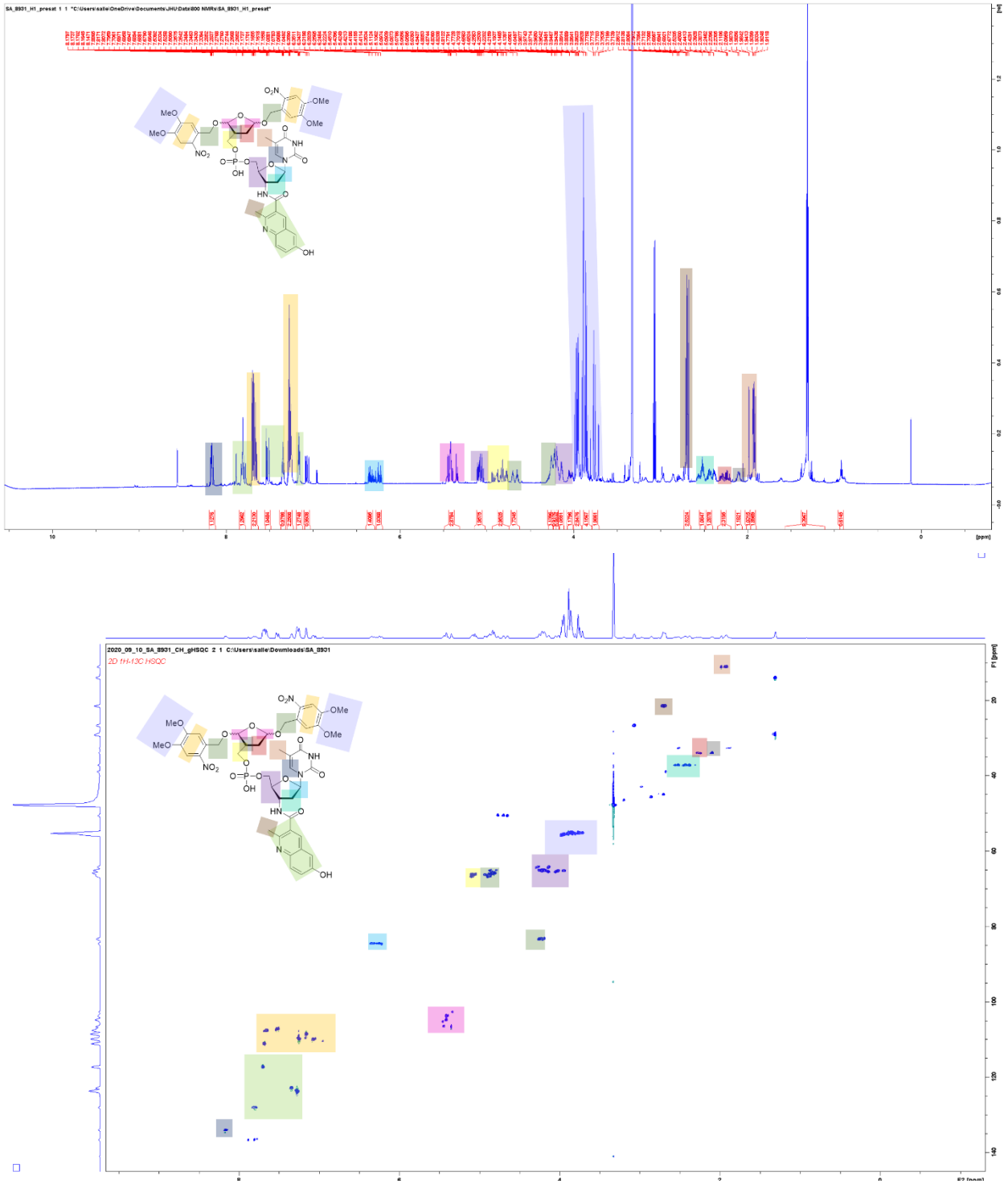


Table 25. Chemical shift analysis of **84** NMR spectra.

Proton			Carbon	
Chemical Shift	**Splitting (Coupling Constant)**	**Integration**	**Chemical Shift**	**Description**

8.17	M	1	134.2	dT (1/1)
7.80	T (*J* = 8 Hz)	1	127.5	8931 (1/4)
7.69	Dd (*J* = 8, 7.2 Hz)	2	107.4	Nv (2/2)
7.52	M	1	117.4	8931 (2/4)
7.35	M	1	122.9	8931 (3/4)
7.27	Quint/d (*J* = 6.6, 0.2 Hz)	2	123.7	Nv (2/2)
7.16	Quint (*J* = 3 Hz)	1	110.0	8931 (4/4)
7.07	Q (*J* = 7.2 Hz)		110.2	
6.35	M	1	84.6	C1’
6.25	M	1		
5.42	M	3	104.8	Acetal (2/2)
				C3’ (1/1)
5.10	M	2	66.4	DOB linker
4.88	M	2	66.4	Nv methylene (2/4)
4.70	M	2	50.7	
4.65	Quint (*J* = 5 Hz)	1	50.7	
4.25	T (*J* = 4 Hz)	1	65.2	C4’ (1/1)
4.22	Q (*J* = 4 Hz)	1	65.5	C5’ (1/2)
4.19	M	2	83.4	Nv methylene (4/4)
4.14	Q (*J* = 5 Hz)	1	65.5	C5’ (2/2)
4.05	M	1		
3.97	Dt (*J* = 14, 5 Hz)	3	55.5	Nv OMe (3/12)
3.87	M	4	55.3	Nv OMe (7/12)
3.77	M	2	55.2	Nv OMe (9/12)
2.70	M	3	21.4	8931 Me (3/3) – **chemical shift in 8931 acid Me is 3.1 ppm**
2.52	M	1	37.1	C2’ (1/2)
2.44	M	1	37.1	C2’ (2/2)
2.23	M	2	33.9	DOB C2’ (2/2)
2.11	M	1	33.9	C3’ methine
1.98	S	1	14.1	dT Me (1/3)
1.93	M	2	14.1	dT Me (3/3)
			26.7	
			32.8	
			39.0	
			39.9	
			41.2	
			65.9	
			66.0	
			66.8	
			83.4	

	104.0	
	104.9	
	106.8	
	107.6	
	110.6	
	123.5	
	127.3	
	130.0	
	139.1	
	142.3	Quaternary Carbons 8931 (5) dT (2 x 1) amide (1) Nv (3 x 1)
	143.1	
	147.7	
	152.5	
	155.9	
	157.5	
	164.9	
	165.0	
	169.7	

Analysis of **89**:

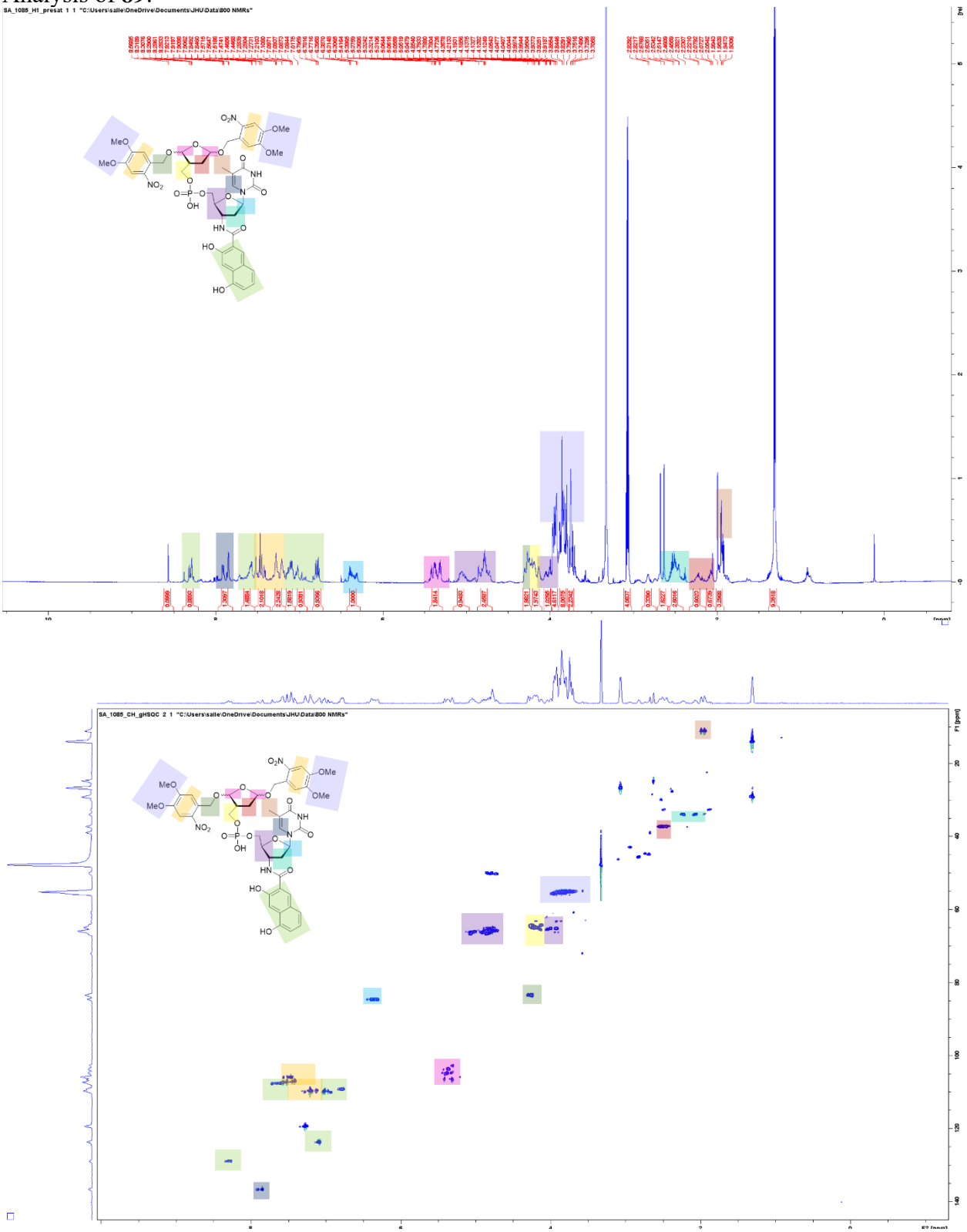


Table 26. Chemical shift analysis of **89** NMR spectra.

Proton			Carbon	
Chemical Shift	**Splitting (Coupling Constant)**	**Integration**	**Chemical Shift**	**Description**
8.31	S	1	129.1	1085 (1/5)
7.84	S	1	136.6	dT aromatic (1/1)
7.59	m	1	107.7	1085 (2/5)
7.46	M	2	105.8	Nv (2/2)
7.28	S	2	110.6	Nv (2/2)
7.09	m	1	128.3	1085 (3/5)
7.01	M	1	109.9	1085 (4/5)
6.78	M	1	109.2	1085 (5/5)
6.38	M	1	84.6	C1' (1/1) (~85 ppm)
5.40	m	2	104.8	Acetal (2/2)
5.06	M	1	66.2	C3' (1/1)
4.78	M	2	65.8	C5' (2/2)
4.28	M	2	83.9	Nv linker *(2/4)*
4.18	M	2	66.3	DOB linker (2/2)
4.02	M	1	65.2	C4' (1/1)
3.93	M	4	55.2	Nv OMe (4/12)
3.83	M	8	55.2	Nv OMe (8/12)
3.75	M	2	55.2	Nv OMe (12/12)
2.51	M	2	37.3	DOB C2' (2/2)
2.23	M	1	33.9	C2' (1/2)
2.07	M	1	33.9	C2' (2/2)
1.95	M	3	11.3	dT Me (3/3)
Missing Nv linker (2/4) Missing DOB methine			14.2	
			24.8	
			26.8	
			29.4	
			39.0	
			55.4	
			55.5	
			107.4	
			110.0	
			128.4	
			129.68	
			129.75	
			139.3	
			147.6	
			150.9	
			151.5	

	153.3	Quaternary Carbons 1085 (5) Amide (1) dT (2 x 1) Nv (3 x 1)
	154.4	
	157.5	
	165.0	
	169.3	

Analysis of Precursor to **93**:

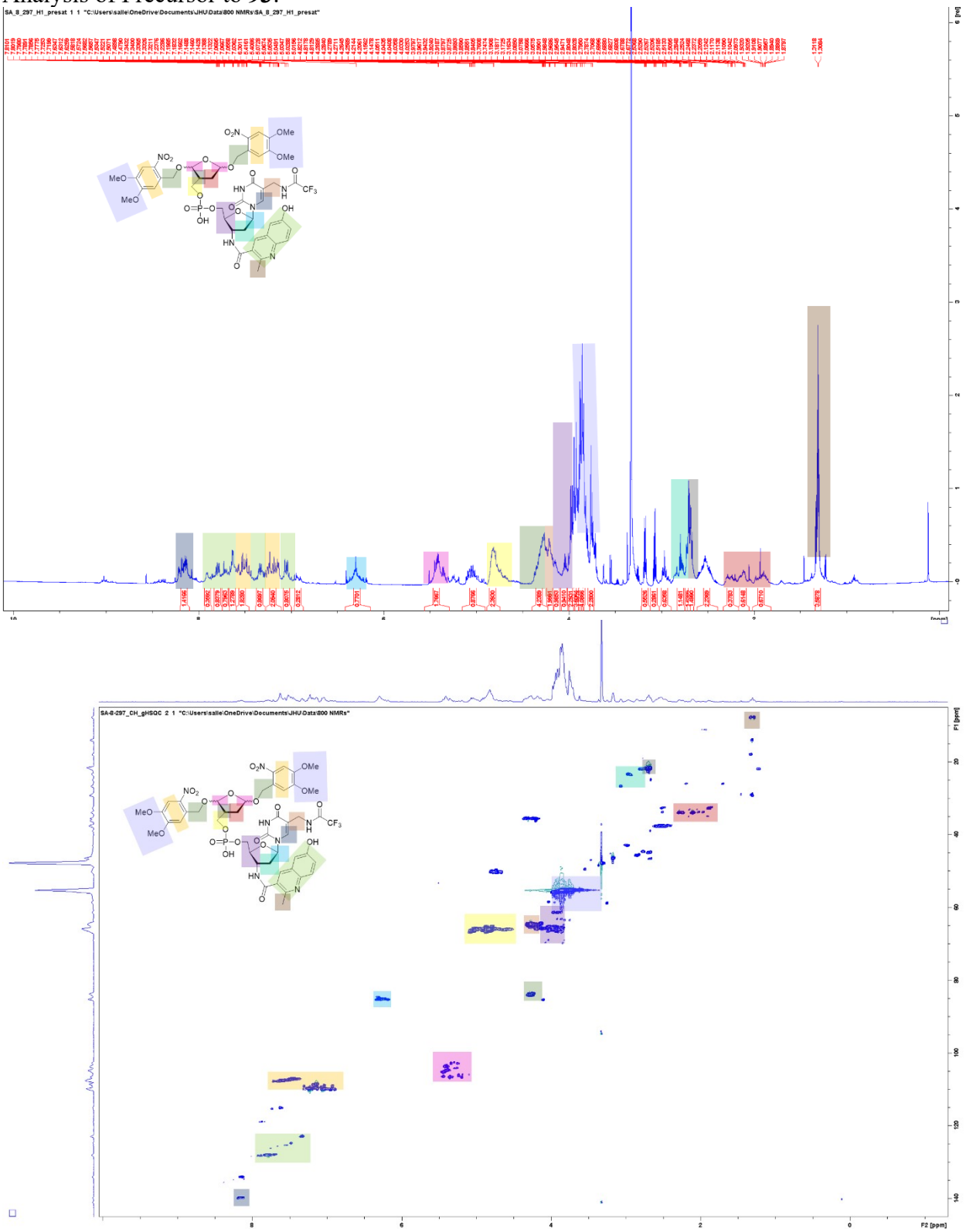


Table 27. Chemical shift analysis of Precursor to **93** NMR spectra.

Proton			Carbon	
Chemical Shift	**Splitting (Coupling Constant)**	**Integration**	**Chemical Shift**	**Description**
8.20	M	1	139.8	dT (1/1)
7.80	M	1	134.2	8931 (1/1)
7.71	M	1	128.2	8931 (1/2)
7.63	D (J = 2.8 Hz)	1	107.6	8931 (3/4)
7.54	M	2	107.4	Nv (2/2)
7.34	D (J = 2.8 Hz)	1	123.0	8931 (4/4)
7.20	S	2	108.5	Nv (2/2)
7.06	M	1	110.0	
6.30	S	1	85.4	C1' (1/1) (~85 ppm)
5.41	M	2	104.8	Acetal (2/2)
5.05	M	1		
4.81	M	2	66.3	DOB linker (2/2)
4.28	M	4	83.7	Nv linker (4/4)
4.21	M	2	64.8	Amide methylene (2/2)
4.14	M	1	66.3	C4' (1/1)
4.05	Q (J = 6 Hz)	1	65.8	C3' (1/1)
3.97	D (J = 14 Hz)	2		C5' (2/2)
3.93	D (J = 14 Hz)	3	55.2	Nv OMe (3/12)
3.87	M	4	55.2	Nv OMe (7/12)
3.85	M	4	55.2	Nv OMe (11/12)
3.76	M	2	55.2	Nv OMe (12/12)
				(1/2)
2.80	T (J = 6 Hz)	1	21.9	C2' (1/2)
2.70	D (J = 6 Hz)	1	21.9	C2' (1/2)
2.67	M	1	23.4	DOB methine
2.52	M	2	37.6	
2.11	m	1	32.8	DOB C2' (1/2)
1.88	M	1	32.8	DOB C2' (2/2)
1.31	M	3	26.7	8931 Me (3/3)
			7.9	
			14.2	
			17.9	
			19.0	
			21.4	
			21.8	
			29.4	
			109.1	
			109.5	

	110.1	
	1154	
	116.9	
	126.2	
	127.3	
	129.8	
	130.0	
	139.1	
	139.2	
	139.7	
	142.3	Quaternary Carbons 8931 (5) Amides (2) dT (2 x 1) Nv (3 x 1)
	147.7	
	150.7	
	151.3	
	152.5	
	153.3	
	153.7	
	155.9	
	156.3	
	157.2	
	157.4	
	157.5	
	163.5	
	169.7	

Curriculum Vitae

Shelby C. Yuhas, Ph.D.

5627 Alhambra Ave Baltimore, MD 21212 | Phone: 908.907.2791 | e-mail: sallen62@jhu.edu

Professional Summary

- Organic and biochemist with 8 years of experience in the development of small molecule compounds with drug-like properties.
- Synthesized and characterized (structurally and biochemically) irreversible DNA repair inhibitors and demonstrated synergistic effects with cancer drugs in cells.
- Contributed to a multidisciplinary research project as the primary graduate student.
- Researched several scientific topics, unrelated to my academic work, for comprehensive research reports.
- Excellent written and oral communication skills from multiple years of experience as a teaching assistant, residential advisor, and other roles in customer service.

Education

Johns Hopkins University (JHU) 2015 - present
M.A. in Chemistry (Aug 2017) (Cum GPA 3.76/4.00)
Ph.D. in Chemistry (Dec 2020)
Thesis Title: Identification of Irreversible Inhibitors of DNA Polymerase β
Thesis Advisor: Professor Marc M. Greenberg

The College of New Jersey (TCNJ) 2011 - 2015
B.S. Chemistry (ACS Certified degree), Marketing Minor (Cum GPA 3.64/4.00)

Research Experience

Graduate Research, Johns Hopkins University, MD 2016 -present

- Designed and synthesized small molecule inhibitor libraries against polymerase beta (Pol β), a key enzyme in DNA repair.
- Characterized all synthesized compounds using structural and analytical techniques (i.e. NMR, MS, IR).
- Experienced in chromatographic techniques (i.e. HPLC, UPLC, FPLC).
- Expressed and purified polymerase enzyme(s) from *E.coli*.
- Identified and validated hit compounds via *in vitro* biochemical assays (e.g. enzyme kinetics, fluorescence anisotropy).
- Knocked down Pol β in mammalian cell lines by RNA interference.
- Characterized inhibitor effects in multiple mammalian cell lines (i.e. HeLa, MCF-7, and MDA-MB-436) using cell viability and clonogenic assays.
- Mentored 2 graduate students.

Undergraduate Research, The College of New Jersey, NJ 2013 - 2015

- Organized a research proposal and initiated a novel project.
- Improved synthesis of benzyl azetidine, a common scaffold in pharmaceutical drugs.
- Synthesized organometallic platinum complexes with anti-cancer properties.

Other Experience

Head Teaching Assistant, Johns Hopkins University 2018

- Oversaw 8 teaching assistants and 225 students.
- Served as substitute lecturer for my advisor on multiple occasions.

Teaching Assistant, Johns Hopkins University 2015 -2017

- Led weekly discussion lectures for 10-15 students.
- Clarified concepts for students during office hours.

Residential Advisor, The College of New Jersey 2013 -2015

- Responsible for mentoring and aiding 45 students.
- Learned to adapt to the needs of my residents and tailor my communication to best serve them.

Desk Assistant, The College of New Jersey 2011 -2013

Publications

Yuhas, S., Greenberg, M. M. Selective (extra- and intracellular) irreversible inhibition of DNA polymerase β. **2020**. (*in preparation)*

Yuhas, S., Greenberg, M. M. Investigation into DNA polymerase β domain-selective inhibitors. **2020**. (*in preparation)*

Yuhas, S., Greenberg, M. M. DNA polymerase β inhibition is synthetic lethal in BRCA1-deficient cells. **2021**. (*in preparation*)

Research and Academic Honors

Owen Scholar Award, Johns Hopkins University 2015 -2016

- Monetary award for an outstanding member of the entering graduate student class.

ACS Undergraduate Award in Organic and Biochemistry, The College of New Jersey 2015

Gamma Sigma Epsilon, The College of New Jersey 2013 -2015

- Elected Vice President for 2014-2015.
- Chapter Sigma Beta- Chemistry Honors Society

National Residence Hall Honorary (NRHH) Member, The College of New Jersey 2014

- For excellence in residential education due to my work as a residential advisor and desk assistant.

Dean's List, The College of New Jersey 2013, 2014

Professional Memberships, Presentations, and Activities

JHU Chemistry Student Safety Committee 2016 - present

- Elected Chair for 2018-2020.

- Identified and solved safety hazards within the chemistry department.
- Integrated safe and environmentally conscious mindset into experimental design and process protocols.

JHU Student Liaison Committee 2016 - present

- Organized department wide community-engaging activities.
- Organized an annual student-invited distinguished investigator seminar visit.

JHU Chemistry Diversity, Networking, and Advancement 2016 - present

Invited seminar presentation, Chemical Biology Interface Program, Johns Hopkins University 2019

Undergraduate Poster Presentation at ACS Meeting in Denver, CO 2015

Member, American Chemical Society 2014 - present